PARADOX LOST: EXPLAINING THE HISPANIC ADULT MORTALITY ADVANTAGE*

ALBERTO PALLONI AND ELIZABETH ARIAS

We tested three competing hypotheses regarding the adult "Hispanic mortality paradox": data artifact, migration, and cultural or social buffering effects. On the basis of a series of parametric hazard models estimated on nine years of mortality follow-up data, our results suggest that the "Hispanic" mortality advantage is a feature found only among foreign-born Mexicans and foreign-born Hispanics other than Cubans or Puerto Ricans. Our analysis suggests that the foreign-born Mexican advantage can be attributed to return migration, or the "salmon-bias" effect. However, we were unable to account for the mortality advantage observed among other foreign-born Hispanics.

Research has shown that Hispanics in the United States experience lower mortality rates in adulthood than do non-Hispanic whites. It has been argued that this phenomenon is a paradox because Hispanics generally have lower socioeconomic status than do non-Hispanic whites. Research on the relationship between socioeconomic status and health and mortality has consistently shown that low socioeconomic status is significantly associated with poor health and mortality outcomes among both non-Hispanic whites and blacks in the United States (Sorlie, Backlund, and Keller 1995). It is, therefore, deemed paradoxical that Hispanics could have better health and mortality profiles than non-Hispanic whites, a population with a more favorable socioeconomic composition.

In this article, we review, critically evaluate, and empirically test various hypotheses that are associated with this paradox. We begin by reviewing key findings in the literature regarding the existence of a paradox and discuss the hypotheses that have been invoked to explain its existence. The three most prevalent of these hypotheses are data artifacts, migration effects, and cultural or social buffering effects. In our study, we used the National Health Interview Survey (NHIS)-Multiple Cause of Death (MCD) data file to test these competing hypotheses. The NHIS-MCD data set is a linked file with nine NHIS surveys (1986–1994) matched, using the National Death Index (NDI), to data on mortality from 1986 to 1997. We further linked the NHIS-MCD data file to data from the 1990 census Summary Tape File 1 (STF1) by state and county geography code and appended contextual-level measures of residential segregation. We then estimated parametric hazard models for the near-decade-long mortality experience of the population aged 35 and older. Available information on ethnicity enabled us to distinguish between non-Hispanic

---

*Alberto Palloni, Center for Demography and Ecology, University of Wisconsin–Madison. Address correspondence to Elizabeth Arias, National Center for Health Statistics (NCHS), 3311 Toledo Road, Room 7330, Hyattsville, MD 20782; E-mail: EArias@cdc.gov. We are grateful to Robert N. Anderson, Jennifer H. Madans, and James A. Weed, of the NCHS, Centers for Disease Control and Prevention; Nan Astone, of the Department of Population Dynamics, Johns Hopkins University; Richard Rogers, of the Department of Sociology, University of Colorado at Boulder; and two anonymous reviewers for their insightful and detailed suggestions. We are also grateful to the participants in colloquia at the University of Chicago, Johns Hopkins University, the Population Council, the University of Pennsylvania, and the University of Washington for their helpful hints and comments. Research for this article was supported by the National Institute on Child Health and Human Development through a core grant (P30HD05876) to the Center for Demography and Ecology, by National Institute on Aging through a core grant (P30AG17266) to the Center for Demography of Health and Aging, and through research grants R01AG16209, ROGAG18016, and R03AG15673. Finally, we thank Bob Hauser for his linguistic contribution.

whites and several Hispanic subgroups, including Cubans, Mexicans, Puerto Ricans, and all other Hispanics combined (hereafter Other Hispanics). The parametric models account for demographic, socioeconomic, and contextual-level determinants of mortality; for interaction effects of theoretical importance; and for unmeasured individual (fixed) frailty.

BACKGROUND

Evidence of the Hispanic Mortality Paradox

Studies based on various data sources, including national and state vital statistics; local surveys; and, most significantly, national linked data files, such as the National Longitudinal Mortality Study (NLMS)[1] and the NHIS-MCD, have found support for the Hispanic mortality paradox. Most studies have reported that after pertinent demographic and socioeconomic characteristics are controlled, the Hispanic population as a whole fares better in adult all-cause mortality than do non-Hispanic whites (Becker et al. 1988; Liao et al. 1998; Markides 1983; Markides and Coreil 1986; Novello, Wise, and Kleinman 1991; Rogers et al. 1996; Rosenwaike 1987; Sorlie et al. 1993). Using the NLMS, Sorlie et al. (1993) found that among adults aged 25 and older, Hispanics are better off than non-Hispanic whites in all-cause and selected cause-specific mortality outcomes even after age and family income are controlled in multivariate hazard models. Two other studies based on the NLMS, but not exclusively focused on the Hispanic population, reported similar results (Singh and Siahpush 2001, 2002). One of these studies (Singh and Siahpush 2001), which used multivariate hazard models estimated for adults aged 25 and older and controlled for demographic and socioeconomic characteristics, found that both Hispanic men and Hispanic women have significantly lower all-cause mortality than do non-Hispanic whites. The other study (Singh and Siahpush 2002) found that U.S.-born Hispanics have lower all-cause mortality than do U.S.-born non-Hispanic whites and that foreign-born Hispanics have lower all-cause mortality than do foreign-born non-Hispanic whites.

Studies that have disaggregated the Hispanic population by national subgroup have reported varying support for the Hispanic mortality paradox. For example, using the NLMS, Abraido-Lanza et al. (1999) found lower hazard ratios for each of four Hispanic subgroups relative to non-Hispanic whites after they accounted for age, education, and family income. After they adjusted for education and family income, they also found lower hazard ratios for U.S.-born Hispanics (excluding Cubans and Puerto Ricans) than for U.S.-born non-Hispanic whites.

LeClere, Rogers, and Peters (1997), who investigated an individual Hispanic subgroup, Mexican Americans, documented similar findings on the basis of an earlier release of the NHIS-MCD (1986–1991) data. Using estimates from multivariate hazard models controlling for a number of demographic and socioeconomic characteristics, they reported that both Mexican American men and Mexican American women have significantly lower all-cause mortality than do non-Hispanic whites.

In contrast to these results, one study, based on the 1989–1995 NHIS-MCD data file, suggested that the mortality of Mexican Americans and Other Hispanics does not differ significantly from the mortality of non-Hispanic whites after age, sex, nativity, education, income, and marital status are controlled (Hummer et al. 1999). However, the authors reported that foreign-born Mexicans and foreign-born Other Hispanics have significantly lower mortality rates than do native-born non-Hispanic whites and that the foreign-born mortality advantage can be detected only among middle-aged adults (aged 45–64) and the elderly (aged 65 and older). Both these findings were replicated in our research.

Using 1986–1995 NHIS-MCD data, Hummer et al. (2000) compared all-cause mortality outcomes of each of five Hispanic subgroups (Mexicans, Puerto Ricans, Cubans,

1. The NLMS consists of a series of Current Population Surveys linked with vital statistics mortality data.

Central and South Americans, and Other Hispanics) with those for non-Hispanic whites. They found that Mexicans, Central and South Americans, and Other Hispanics have significantly lower mortality than do non-Hispanic whites and confirmed that these differences are found mostly in the older age group (aged 65 and older) for both men and women. However, they found no significant difference in the mortality rates of Puerto Ricans, Cubans, and non-Hispanic whites.

In summary, the findings of previous studies offer general support for the existence of a Hispanic adult mortality advantage. When grouped, Hispanic adults exhibit lower mortality rates than do non-Hispanic whites even after pertinent demographic and socioeconomic characteristics are taken into account. An important revelation of this review is that findings regarding the Hispanic mortality advantage have been fairly uniform across two distinct data sets, the NLMS and NHIS-MCD. This revelation is significant because findings that are based on linked data sets are not affected by the problem of ethnic misidentification, a shortcoming that is inherent in vital statistics.

Despite these advances, however, the literature has left open several questions: Does the adult mortality advantage apply to all Hispanic subgroups or only to a few of them? If so, which ones enjoy a more favorable status? Are differences comparable across gender, or are they visible only among men? And, finally, is the advantage confined to a few age groups, or does it permeate the entire age pattern of adult mortality of these groups?

Explanations of the Hispanic Health and Mortality Advantage

There are three standard explanations for the observed mortality advantage of Hispanics. The first explanation suggests that the Hispanic advantage is an illusion produced by data artifacts. The second is based on the idea that there are conditions associated with migration into and out of the United States that could favor the health profile of the resident migrant population. The third rests on the argument that the Hispanic population enjoys conditions that lead to “cultural or social buffering effects” that imply different behavioral profiles, secure emotional support from social networks, and enhanced self-control and self-efficacy. The second explanation applies only to migrants (non-U.S.-born Hispanics), whereas the remaining two may be applicable to all Hispanic groups. If any of the mechanisms on which these three explanations rely operate, the main outcome will be that after measurable determinants are controlled, Hispanic mortality will appear to be lower than mortality of individuals who belong to other ethnic groups and who share other relevant traits.

Data artifacts. There are three equally salient data problems that may lead to the appearance of a Hispanic mortality advantage. The first two are shared by all data sets for which mortality rates (or the prevalence of diseases) require ethnic self-identification and self-reporting of ages. The third is pertinent only for data sets, such as the one we used in our study, in which rates are calculated by matching death and survey records.

Ethnic identification. The underreporting of Hispanic origin on U.S. death certificates, which are the prime data-collection tool for mortality statistics, is a significant problem. Mortality rates in the United States are based on two distinct data sources: the vital statistics system (numerator) and the census population enumeration (denominator). Ethnic identification in the denominator is usually and mainly self-identification. Ethnicity in the numerator is reported by someone other than the decedent. Incongruence between the classification of Hispanic origin in numerators and denominators leads to artificially low death rates for Hispanics relative to non-Hispanics and to the appearance of an advantage. On the basis of analyses of the NLMS, it has been estimated that about 7% of Hispanics are not recorded as Hispanic on the death certificates (Rosenberg et al. 1999).

Misreporting of ages. The second source of data artifact is associated with the misreporting of ages. It has been shown that some populations in Latin America (Dechter and Preston 1991) and some Hispanic subgroups in the United States (Rosenwaike 1991;

Rosenwaike and Preston 1983) tend to overstate their ages, particularly those who are older than age 55 or 60. The net overstatement of ages in the population will depress mortality rates at older ages, thus producing the illusion of a more-benign mortality pattern. A similar outcome occurs when the overstatement of ages affects the age distribution of deaths. When both distortions are present, errors will be offset, but typically, the net effect will be to bias the death rates downward.

Mismatches of records. The third source of error that could lead to downward biases in the mortality rates of some ethnic groups applies only to data sets in which mortality rates are constructed by matching deaths that occurred during an interval of time to populations that were enumerated at the beginning of the interval. The success rate of the matching procedure that links a death record to a population record is variable across individuals and groups. Matching rates are usually associated with missing or erroneous information in some of the key characteristics that make the matching algorithm feasible, such as surnames or social security numbers. Although no good sources have documented differential matching rates by ethnic groups, it is suspected that these rates could be lower in populations whose identification via universal identifiers is more difficult to obtain because of their legal status. The result is to impart a downward bias to mortality rates of ethnic groups that are more heavily composed of individuals whose legal status is questionable and whose identifiers are less complete or less reliably recorded.

Migration effects. Two hypotheses are part of the explanation based on migration effects: the "healthy-migrant effect" and the "salmon-bias effect." If the observed difference between Hispanic and non-Hispanic mortality is a result of migration effects, one cannot conclude that there are characteristics—genetic, socially produced, or culturally acquired—conferred upon individuals by virtue of their membership in the group, that translate into health advantages and lower mortality. One cannot do so because the observed difference between migrants and nonmigrants is net of measured characteristics, which almost certainly will not include all those that are relevant to both migration decision making and to health and mortality.

The healthy-migrant effect posits that the selection of healthy migrants to the United States accounts for the epidemiological paradox (Abraido-Lanza et al. 1999; Palloni and Morenoff 2001; Sorlie et al. 1993). Hispanic migrants are selected from the origin population for certain traits, including better physical and psychological health. The population of successful migrants is not a random draw from the health distribution of the origin population. On average, migrants are healthier than those who do not migrate and may be healthier than the average individual in the receiving population.

The salmon-bias effect is due to a phenomenon experienced by some non-U.S.-born Hispanic subgroups—the propensity to return to the country of origin following a period of temporary unemployment and/or illness (Abraido-Lanza et al. 1999). Return migration will result in artificially lower mortality rates for the Hispanic population for one of two reasons. First, death rates are typically calculated for a period using denominators that were estimated from a baseline population and a numerator that includes all relevant events observed during the period. If return migration occurs, a downward bias in the estimated rates will result, irrespective of whether return migrants are less healthy than those who stay (Abraido-Lanza et al. 1999). If there is no significant difference between those who stay and those who leave, the rates can be corrected by excluding all individuals who are known to have left the United States from the initial exposure counts. We refer to this effect as a Type 1 "return-migrant effect or bias."

Second, to the extent that returning migrants are more likely to be in poor health and exposed to higher risks of mortality than are those who stay, the death rates for a given period will be biased downward *even if one were able to adjust denominators by excluding those who left the country*. We refer to this effect as a Type 2 "return-migrant effect or bias."

Because it provides an identifying clue, the magnitude of the impact associated with healthy-migrant and return-migrant effects should, at least in theory, vary by age. The return-migrant effect (especially Type 2) is likely to have greater salience at older ages, when increases in morbidity augment the population who is at risk of experiencing return migration. Return-migration effects should lead to mortality rates at older ages that are too low.

On the other hand, the healthy-migrant effect should be more visible at younger ages, specifically within the age interval over which the age distribution of *recent* migrants is largely concentrated. It should be so because as both migrant and domestic populations age, their composition by frailty tends to converge, so that at very old ages, all healthy-migrant effects should vanish (Palloni and Morenoff 2001).

Cultural effect. The main premise of the "cultural" hypothesis is that Hispanics' mortality advantage is a function of social and cultural characteristics that differentiate this population from the non-Hispanic population, rather than selection (Abraido-Lanza et al. 1999; LeClere et al. 1997; Sorlie et al. 1993). It proposes that culture affects mortality outcomes by influencing individual health and lifestyle behaviors, family structure, and social networks.

First, culture of origin may shape the behavioral-risk profile of individuals. For example, diet is closely tied to cultural practices, as is the prevalence of smoking, alcohol consumption, habitual exercise, and the use of preventive medical care. A second type of effect shapes the nature of the social environment of individuals and operates through norms and beliefs about family relationships and obligations. Such norms and beliefs may influence the propensity to live alone or in extended families, the density of social networks, the amount of social support exchanged, and the sense of control and self-efficacy (Arias 1998). It is suspected (although not conclusively proved) that health status and mortality are related to individuals' ability to participate actively in social networks, to establish bonds of reciprocal obligations through which they derive emotional and material support, and to enhance their sense of control (Mendes de Leon and Glass 2002). This is a plausible explanation in that there may be some physiological benefits in the form of the dilution of "allostatic" loads that accrue to social integration (Adler and Ostrove 1999). One would expect that individuals who are more successful in establishing an identity and forging strong social and cultural ties will be exposed, ceteris paribus, to conditions that are more conducive to good health and a low risk of mortality. By contrast, those who fail to establish social linkages may be left in a disadvantaged position.

It follows that a successful accounting of the Hispanic mortality paradox using the cultural explanation must verify the joint occurrence of the following three regularities: (1) other things being equal, Hispanics who share advantageous mortality and health conditions must also share either beneficial behavioral-risk profiles and/or denser social networks and social, emotional, and material support than must individuals who do not display the advantage; (2) Hispanics who are not well-integrated into social networks and who receive less social support will experience higher exposure to health and mortality risks and will not share the advantage from which other members of the same ethnic group benefit; and (3) the mortality advantage should fade with increasing assimilation into the receiving country if the latter implies either the acquisition of a less-healthy behavioral profile or the abandonment of norms and behaviors that secure social support.[2]

2. The data set we used did not include information on behaviors such as exercise, smoking habits, and alcohol consumption. Therefore, we did not test the behavioral-profile interpretation of the cultural explanation.

METHOD

Data Source

The NHIS, conducted annually in the United States since 1957, is a nationally representative multistage probability sample of the civilian noninstitutionalized American population. The 1985–1994 design obtained information from approximately 49,000 households, including about 132,000 individuals per year, through personal interviews. The annual response rate was over 95%. The 1985–1994 NHIS design consisted of two basic parts: (1) a core questionnaire containing basic demographic and health questions and (2) one or more modules with questions related to current health topics. The basic or core questionnaire was repeated yearly. Questions included those on demographic, socioeconomic, and health-status characteristics, such as age, sex, race, education, family income, and self-assessed health status. The core questionnaire also included a set of questions about disability, visits to physicians, chronic conditions, and hospital stays. Special modules with questions on specific health topics changed yearly and included such topics as alcohol use, smoking, health care, and health insurance (National Center for Health Statistics, NCHS, 1989).

Beginning in 1986, linkage information for NHIS respondents aged 18 and older was collected to match the NHIS individual records with those of other data systems, including the NDI. The NHIS record is linked to the NDI on the basis of a series of combinations of 12 identifiers, including social security number, first and last names, father's surname, and month and year of birth. Once a linkage is made and vital status is ascertained, the records are linked to the national vital statistics data on multiple causes of death, resulting in the NHIS-MCD file. Information on mortality is available for the NHIS surveys beginning with 1986 and ending with 1994 for deaths through December 31, 1997 (NCHS 2000).

In addition to vital status, state- and county-level information from the 1990 census STF1 was linked to the NHIS for 1986 to 1994. A measure of Hispanic geographic concentration, the Isolation Index, was estimated from census data for use as a proxy for cultural effects.[3] To facilitate the analyses, we used a 10% random sample of the non-Hispanic white population in combination with the full sample of the Hispanic subgroups. Because survey years 1986–1988 did not include a variable for nativity or duration of residence in the United States, an important variable in our analyses, we used the 1989–1994 NHIS with 1989–1997 mortality follow-up. The population exposed and the frequencies of relevant events for selected subgroups are listed in Appendix B.

Basic Model and Enhancements

The data set enabled us to estimate adult mortality over a nine-year period for a population aged 35 and older at the baseline. Because of the design of the study, we could assess only the effects on individual mortality risks of characteristics elicited at the outset and could not evaluate the effects of changes in these characteristics. We estimated standard parametric hazard models to assess the existence and magnitude of a mortality advantage among both foreign-born and U.S.-born Hispanics and to identify the mechanisms that may produce it. Both tasks are complicated because data artifacts and various mechanisms that genuinely produce an advantage can lead to similar observable patterns. Throughout, we searched for identifying signals that helped us separate the contribution of artifacts and other mechanisms.

3. Given the lack of individual data on the strength of social networks, we used Hispanic concentration at the neighborhood level on the assumption that cultural or social buffering effects can occur only in areas with high concentrations of coethnics. See Appendix A for a description of the Isolation Index, as well as all other variables that were used in this study.

We used a standard parametric hazard model to estimate effects on mortality for the decade-long follow-up period for individuals aged 35 and older at the time of the baseline survey.[4] We assumed that a Gompertz model represents well the profile of an increase in mortality for ages 35 and older. Throughout, we rescaled age to be the difference between age at the onset of the study and 35, so that Gompertz's constant refers to the mortality rate at age 35. With this modification, the hazard rate t years into the study can be expressed as follows:

$$\mu_i\,(t \mid X_i\,;\,\mathbf{Z_i}) = \mu_o\,(X_i - 35 + t)\,\exp(\boldsymbol{\beta}\mathbf{Z_i}) = \alpha\,\exp(\gamma\,(X_i - 35 + t))\,\exp(\boldsymbol{\beta}\mathbf{Z_i}), \tag{1a}$$

or,

$$\mu_i\,(t \mid X_i\,;\,\mathbf{Z_i}) = \alpha\,\exp(\gamma\,t)\,\exp(\gamma\,(X_i - 35))\,\exp(\boldsymbol{\beta}\mathbf{Z_i}), \tag{1b}$$

where $\mu_i\,(t \mid X_i\,;\,\mathbf{Z_i})$ is the hazard rate t years after the onset of the study for an individual i aged X_i at the outset who is characterized by a vector of attributes $\mathbf{Z_i}$. The expression α $\exp(\gamma\,(X_i - 35 + t))$ in Eq. (1a) refers to the standard Gompertz mortality rate evaluated at age $X_i + t$; $\boldsymbol{\beta}$ is a vector of effect parameters; α is Gompertz's constant scaled to represent the mortality rate at age 35; and γ is an ancillary parameter, the slope of the hazard rates above age 35.

It is clear from Eq. (1b) that the parameter for age (rescaled) need not equal γ. It is only because Eq. (1b) is a reexpression of Eq. (1a) that this equality is necessary. In fact, one could think of Eq. (1b) as the constrained version of the more general model, Eq. (1c):

$$\mu_i\,(t \mid X_i\,;\,\mathbf{Z}_i\,) = \alpha\,\exp(\gamma\,t)\,\exp(\delta\,(X_i - 35))\,\exp(\boldsymbol{\beta}\mathbf{Z_i}). \tag{1c}$$

Although Eq. (1c) is plausible, only Eq. (1b) is compatible with the Gompertz model in Eq. (1a). This finding suggests a test to validate Eq. (1a): if an estimated model that constrains δ to equal γ does not fit the data as well as a model that leaves the parameters unconstrained, we have prima facie evidence suggesting either that the underlying hazard cannot be reproduced by the Gompertz model in Eq. (1a) or that the effects of age and/or duration in the study are biased because of measurement errors. We return to this issue in the Results section.

In addition to verifying that the constrained form of the model (Eq. (1b)) is acceptable, we must ensure that the estimated values of α and γ are within the expected range: α must be close to the observed mortality rate in the nonblack population (non-Hispanic whites and Hispanics) at age 35, and γ must fall within the known range for a population, such as that of the entire United States (between .06 and .12).

To test some of our hypotheses, the vector **Z** includes variables that reflect ethnic group, marital status, socioeconomic characteristics, nativity and duration of residence, state of residence, and a constructed index of ethnic isolation. When required, we included suitable interaction terms. Appendices A–C contain a full description of the variables, as well as descriptive statistics of the sample.

Finally, we investigated the healthy-migrant and return-migration effects and identified of data artifacts by generalizing Eq. (1a) in three different ways: by assuming that the effects of age are not invariant over the age span, by formulating a model in which the slope is a function of covariates, and by posing the existence of unmeasured frailty (unmeasured heterogeneity).

---

4. Our initial intention was to study the mortality experience at ages older than 40. Because individuals who were younger than 40 at the baseline would contribute variable amounts of exposure at ages 40 and older during the follow-up period, we opted for a compromise solution and included individuals who were aged 35 and older at the baseline.

Summary of Key Hypotheses

In the following sections, we present our tests of evidence to verify or reject the competing hypotheses that we formulated earlier. For the sake of clarity, we summarize these hypotheses into four groups.

Data artifact (Hypothesis 1). The Hispanic adult mortality advantage[5] is an illusion; it is merely the consequence of data artifacts that affect ethnic identification, age misreporting, and/or differential rates of mismatching of death and population records.

Healthy-migrant effect (Hypothesis 2). The Hispanic advantage is the result of migration selection effects, whereby migrants who enter and settle in the United States are disproportionately drawn from groups at origin whose health status is above average.

Salmon-bias effect (Hypothesis 3). The Hispanic advantage is a result of the return migration of non-U.S.-born Hispanics who return to their countries of origin when they are ill.

Cultural effect (Hypothesis 4). The Hispanic advantage is the outcome of "cultural capital," or characteristics that are associated with culturally defined behaviors and/or Hispanic social networks that reinforce the intensity and magnitude of social support. As we mentioned in footnote 2, we tested only the second part of this hypothesis.

RESULTS

Mortality Profiles

We first examine observed mortality patterns at ages 35 and older by gender. To avoid cluttering of figures, we refer the reader to a companion publication (Palloni and Arias 2003) and only summarize the main findings here. Overall, Puerto Ricans and Cubans exhibit mortality rates that are slightly higher than those of non-Hispanic whites, whereas the rates for Mexicans are virtually identical to those of non-Hispanic whites and the rates for Other Hispanics are considerably lower, especially at younger ages. Disparities between U.S.-born and foreign-born Mexicans and Other Hispanics are significant. The differences for Puerto Ricans[6] and Cubans (U.S.- and non-U.S. born) are unimportant. Foreign-born Mexicans have considerably lower mortality rates than do both U.S.-born Mexicans and non-Hispanic whites. The differences are particularly salient at older ages, as would be expected if Type 2 biases (or the overstatement of ages) were of some importance. Both foreign-born and U.S.-born Other Hispanics also have lower mortality rates than do non-Hispanic whites, but the differences are especially large among the foreign born at any age.

In summary, on first blush, the Hispanic adult mortality advantage is not uniform across all Hispanic groups. It is especially prominent among Other Hispanics and less so among Mexicans. Furthermore, it is much more salient among foreign-born Mexicans and Other Hispanics than among those who were born in the United States. Thus, Hispanic group and nativity are both important characteristics and need to be considered explicitly in all models.

Models for Males and Females

Exploratory analyses with an array of simple models (results not shown; see Palloni and Arias 2003) suggested some important patterns. First, there is strong evidence to

5. Throughout, we use the term *Hispanic adult mortality advantage* (or simply *Hispanic advantage*) to mean the difference in mortality rates between Hispanics and non-Hispanic whites aged 35 and older.

6. Puerto Ricans are U.S. citizens, whether they were born on the U.S. mainland or on the island of Puerto Rico. In this study, the term *foreign-born Puerto Ricans* refers to Puerto Ricans who were born on the island of Puerto Rico.

conclude that nativity effects are indeed important and that these effects are specific to each Hispanic group.

Second, according to the baseline model (which includes only baseline hazard parameters), mortality rates in the 35–39 age group are .0017 and .0012 for men and women, respectively. These values are remarkably consistent with those that were observed in the U.S. national life tables for white men and women (.0022 and .0010, respectively) for 1991 (a year that was about midway through the follow-up period).

Third, the slope parameter for men and women (~.086 in the null model) falls nicely in the middle of the range for a population like the United States. In addition, in a model for men that does not constrain the effects of the variable delta = (X_i – 35) to be identical to γ, the Gompertz slope parameter performs just as well as in a model that imposes the constraint. This is prima facie evidence that the functional form imposed on the data captures satisfactorily the age effects on the hazards. The same is only approximately true for the female sample, in which the fit is slightly better when the age effects and slope are not constrained to be identical. All other estimates, including the slope, are insensitive to the specification of the constraint on the parameter for δ. As we discuss later, the fact that the constrained and unconstrained models perform equally well not only suggests that the Gompertz formulation is reasonable but also helps to rule out partially the possibility that estimates of ethnicity are heavily contaminated by the overstatement of ages.

The fourth regularity is that the Hispanic adult mortality advantage is a trait that is found only among Mexicans and Other Hispanics, not among Cubans or Puerto Ricans. The advantage is characteristic of foreign-born Mexicans and Other Hispanics, not of U.S.-born individuals. For men, the overall mortality rate among non-U.S.-born Mexicans is approximately equal to exp(–.26), or 77% as high as the mortality rate associated with non-Hispanic whites, whereas the rate for Other Hispanics is even lower, about exp(–.61), or 54% as high. Both these effects are statistically significant. In contrast, the mortality rates of Puerto Ricans, Cubans, and either U.S.-born Mexicans or U.S.-born Other Hispanics are not significantly different from those of non-Hispanic whites. The same patterns apply to women with one exception: the advantage of foreign-born Mexican women is of a lower magnitude and is statistically insignificant (the *t* statistic is about 1.3). Without exception, the numerical value of the relative advantage is much larger among Other Hispanics than among Mexicans.

Table 1 displays key statistics for fitted models in the male and female samples. These are additive models following Eqs. (1a) and (1b) and covering a broad range of specifications, all including controls for demographic and socioeconomic indicators that reflect conditions that are known to influence adult mortality. In light of the diagnostics made before, we used refined variables for ethnicity that enabled us to discriminate between Hispanics by nativity. We first estimated the least-parsimonious model, one in which the effects of ethnicity-nativity groups are left free. In all cases, the non-Hispanic white population is our residual or contrast group.

Arguably, it is only in models that control for relevant compositional factors, such as those in Table 1, in which we must search for an advantage, identify the groups affected, and assess its approximate magnitude. The results of these models are in Model 1 (columns 1 and 3) of Table 1 for men and women, respectively. The log-likelihood-ratio test statistics comparing this model with its baseline equivalent, a model in which only ethnicity and nativity are controlled (see the bottom of Table 1, columns 1 and 3), suggest a much better fit for both men and women when relevant socioeconomic and demographic factors are taken into account.

These estimates reveal a picture of remarkable consistency and regularity that can be summarized with four statements:

Invariance. Introducing marital status or any other factor that is designed to control for socioeconomic conditions, such as education, income, or employment, does not alter

Table 1. Models With Controls for Demographic and Socioeconomic Conditions, Men and Women Aged 35 and Older (*SE* in parentheses)

	Men		Women	
Parameter	Model 1	Model 2[a]	Model 1	Model 2[a]
Baseline Hazard				
Constant	–6.07 (.044)	–6.07 (.039)	–6.88 (.087)	–6.905 (.085)
Delta	.067 (.003)	.067 (.008)	.073 (.0025)	.073 (.0022)
Gamma[b]	—	—	—	—
Ethnicity-Nativity				
Non-Hispanic white	—	—	—	—
U.S.-born Puerto Rican	–.267 (.243)	–.098 (.039)*	.184 (.348)	–.027 (.047)
Foreign-born Puerto Rican	–.106 (.104)	—	–.058 (.114)	—
U.S.-born Cuban	–.365 (.508)	—	.828 (.430)	—
Foreign-born Cuban	–.049 (.117)	—	–.127 (.075)	—
U.S.-born Mexican	–.082 (.074)	—	–.122 (.084)	—
Foreign-born Mexican	–.432 (.034)**	–.433 (.036)**	–.242 (.126)*	–.232 (.127)*
U.S.-born Other Hispanic	–.125 (.153)	—	.050 (.139)	—
Foreign-born Other Hispanic	–.629 (.105)**	–.629 (.104)**	–.730 (.200)**	–.724 (.148)**
Demographic/Economic				
Married	—	—	—	—
Unmarried	.219 (.032)**	.216 (.033)**	.171 (.062)*	.180 (.056)**
< High School	—	—	—	—
High school	–.273 (.069)**	–.275 (.072)**	–.037 (.080)	–.019 (.083)
> High school	–.247 (.080)*	–.249 (.083)*	–.091 (.070)	–.064 (.078)
Family income, first quartile	—	—	—	—
Family income, second quartile	–.074 (.039)*	–.073 (.038)*	–.058 (.090)	–.059 (.092)
Family income, third quartile	–.132 (.039)**	–.133 (.037)**	–.089 (.093)	–.090 (.088)
Family income, fourth quartile	–.255 (.092)*	–.255 (.035)**	–.270 (.077)**	–.273 (.076)**
Employed	—	—	—	—
Unemployed	.359 (.293)	.362 (.301)	.046 (.440)	.038 (.464)
Out of the labor force	.565 (.101)**	.563 (.099)**	.468 (.053)**	.461 (.047)**

(continued)

the advantage revealed before. That is, the Hispanic adult mortality advantage is shared by foreign-born Mexicans and Other Hispanics, but by nobody else. If anything, the magnitude of the advantage is slightly increased in models with extensive controls. In fact, according to Model 1, foreign-born Mexican men have mortality rates that are exp(–.43),

(Table 1, continued)

	Men		Women	
Parameter	Model 1	Model 2[a]	Model 1	Model 2[a]
Sample Size	17,825		21,189	
Events	1,632		1,590	
Log-Likelihood	–5,564.4	–5,565.0	–5,661.9	–5,668.6
Diff. *df*[c]	8	5	8	5
Chi-square[d]	142.0	1.2	62.4	13.4
Upper Tail[e]	< .000	< .94	< .000	< .03

[a]Model 2 constrains the effects of Puerto Rican, Cuban, U.S.-born Mexican, and U.S.-born Other Hispanic to be the same.

[b]When not shown, the parameter for gamma is constrained to be equal to the parameter for delta.

[c]Diff. *df* is the difference in the degrees of freedom between Model *j* and Model (*j* – 1) displayed in the preceding column. In the case of Model 1, the contrast is with a baseline model, controlling only for ethnicity and nativity (see Palloni and Arias 2003 for greater detail). In the case of Model 2, the contrast is with Model 1, and the difference in *df* refers to the number of parameters of Model 1 that are constrained in Model 2.

[d]Chi-square is the chi-square statistic or –2 × (LLc – LLu), where LLc and LLu are the log-likelihood of the constrained and unconstrained models. The contrast for Model 1 is a baseline model, controlling only for ethnicity and nativity, as the constrained model. The contrast for Model 2 uses Model 1 as the unconstrained model.

[e]The upper-tail probability is the cumulated probability above the observed value of the chi-square statistic.

*$p < .0114$ or the absolute value of the t statistic is larger than 2 but smaller than 3.5; **$p < .000116$, or the absolute value of the t statistic is larger than 3.5.

or 65% as large as those of non-Hispanic white men, whereas foreign-born Other Hispanic men have mortality rates that are exp(–.63), or 53% as high.[7]

Regularity. All factors introduced as controls behave as expected, are properly signed, and exert influences that are statistically significant. Thus, for example, the impact of marital status is powerful and, as expected, indicates that single persons have considerably higher mortality rates than do married persons. The same applies to education, income, and employment.

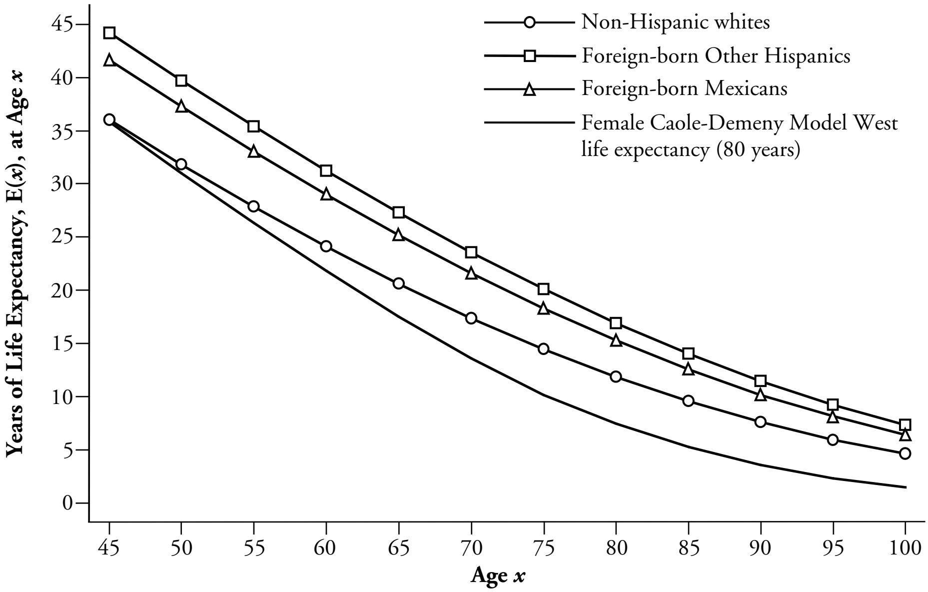
Minimal gender contrasts. The pattern of relationships for women is not identical to that of men, but it is remarkably similar. There are two features that make the pattern for women distinct. One is the lack of strength of the advantage among foreign-born Mexican women. In fact, although the estimated effect is negative, the corresponding t statistic is only 1.92, somewhat below the threshold value of 2.0 that we used to allocate statistical relevance. The other feature is that there is a positive effect on mortality for U.S.-born Cuban women. But, again, the estimated parameter does not quite reach a threshold of statistical significance. All other observed patterns are concordant with those found among men.

General invariance of constrained and unconstrained models. Tests contrasting models in which the effects of age and slope are constrained to be equal to each other with models in which the constraint is absent indicated that the constrained model (and thus the Gompertz baseline) is preferable. Even the exception among models without controls (for women) disappeared.[8]

7. The figures in models with no demographic or socioeconomic controls are 77% and 54%, respectively.

8. The chi-square statistics for contrasts associated with Models 1 and 2 for men and women in Table 1 are not reported there, but their values are too small to reject the null hypotheses that the slope and age effects are identical.

Figure 1. Predicted Residual Life Expectancies, E(*x*)



Statistically significant differences constitute important raw materials for testing theories, but may have remarkably little influence in the lives of actual individuals. What does the advantage detected in these models really amount to? To offer an easily interpretable metric, we translated the differences in mortality rates between foreign-born Mexicans and Other Hispanics, on the one hand, and non-Hispanic whites, on the other, into differences in life expectancies. Figure 1 displays the predicted life expectancies at ages 45 and older for men who, at the onset of the study, were 45 years old. The apparently innocuous effect of –.63 for foreign-born Other Hispanics translates into a difference in life expectancy at age 45 of about eight (7.31) years. For foreign-born Mexicans, the estimated effect of –.43 translates into a difference in life expectancy of about 5 (4.74) years. For women (figure not shown), the differences are slightly lower. Because the life expectancy of white men at age 45 is roughly 39.7, the estimated relative risks translate into an advantage in years of life expectancy at age 45 amounting to 18.5% and 11.9% for foreign-born Other Hispanics and Mexicans, respectively. For women, the relative advantage is on the order of 16% and 9%, respectively.[9]

Are these differences plausible? To place these contrasts in perspective, we also plotted the residual life expectancies in the Coale-Demeny life tables with the highest life

9. Under a Gompertz function, there is an approximate correspondence between relative risks and life expectancy at age 45 that can be expressed as follows: $\ln(E(45)) \sim 3.92 - .99 \times RR$, where RR is the relative risk and E(45) is the life expectancy at age 45. When the estimated coefficient of a 0/1 variable is 0, the RR will be 1 and the value of E(45) about 18.7. When the estimated effect of a 0/1 variable is –.60, the associated RR is .55 and E(45) is roughly 29.23, or about 10 years more than the baseline group. These are all approximations that are based on a range of values of mortality rates at age 35 not exceeding .0050 but not lower than .0010.

expectancy at birth (80 years) in the life tables for women (see Figure 1). This also happens to have been the life expectancy at birth among women in the United States in 2000 (Arias 2002). Note that against this extreme standard, the advantage for foreign-born Mexicans and Other Hispanics is considerable. The implied differences in life expectancy at age 80, for example, are 7.5 to 9.3 years, respectively. These are nontrivial differences, even when considered against the backdrop of a contrast between the non-Hispanic white population and the life table for women (Model West) of about 4.5 years. Are these differences too large and perhaps the product of data artifacts? We now discuss the role that data artifacts may play in our estimates.

Assessing Hypothesis I (Data Artifacts)

The most obvious culprit behind excessively large life expectancies is an across-the-board downward bias in the estimated mortality rates. As we argued earlier, there are three potential sources for this type of bias. We examine each in turn.

Ethnic misidentification. It is unlikely that our data set is affected by this source of error because the very nature of our data set rules out numerator/denominator biases that are associated with inconsistent ethnic identification. Ethnic categorization is derived from the baseline NHIS *self-identification*, rather than from the death certificate's proxy response, in which the bulk of errors and inconsistency are believed to be rooted.[10]

Overstatement of ages. The overstatement of the ages of the baseline population would, indeed, lead to downward biases of mortality rates and to the mirage of a Hispanic adult mortality advantage. However, there are two reasons why this explanation is also unlikely to hold. The first is that the pattern of age overstatement has been found among Hispanic populations in general (Dechter and Preston 1991) and in other ethnic groups in the United States (Preston et al. 2003; Rosenwaike and Preston 1983). But the evidence is somewhat elusive regarding the presence of the bias among some Hispanic subgroups in the United States (Rosenwaike 1991). On the other hand, if the overstatement of ages was, indeed, the culprit and a generalized characteristic of Latin American origin, why is it that we did not find an advantage among other Hispanic groups?

The second reason that undermines this explanation is more technical and relates to the constrained estimation of the coefficients of age and Gompertz's slope. If there was an extensive overstatement of ages, there should be a downward bias on rates, inducing proportionately higher errors at older ages. Overstatement of ages must be reflected in estimated effects of age that do not mirror the passage of time, so that the effect on mortality of a one-year difference in age is less than the effect of a one-year difference in exposure to mortality (controlling for age). In other words, the overstatement of ages leads to an estimate of the coefficient of age that is lower than the estimate of the slope, thus invalidating the use of a (constrained) Gompertz function. But we have shown that, with a singular exception, and then in a marginal way, the constrained version of the model fits as well as the unconstrained one. And the exception is the model for women, an odd finding, since the overstatement of ages in the Hispanic population is seemingly more serious among men than among women.

Mismatching records. A different but related source of error may originate from imperfect, incomplete, or impossible matches. The NHIS-NDI matching algorithm has

10. However, a bias could well exist if misclassification is more likely among those who are in worse health or have higher risk of death. Though possible, this scenario is admittedly somewhat farfetched. We have no knowledge of evidence or even a speculation suggesting that ethnic self-reporting is a function of health status and are not aware of any data that support this the speculation (Sandefur, Campbell, and Eggerling-Boeck 2002). If these errors do exist, they would deflate Hispanics' mortality rates by no more than 7%. But even reducing the size of the estimated regression coefficient to adjust for this possibility leads to a large and statistically significant estimate of effects.

been shown to be highly reliable, with an overall 98% of successful matches based on analyses of known decedents from the active follow-up of the National Health and Nutrition Examination Survey I cohort. However, this cohort contained few Hispanics, and there is no way to assess the quality of the algorithm for this population. As a result, the possibility of errors that are due to erroneous matches or mismatches cannot be ignored (NCHS 2000).

Mismatching death and survey records can lead to downwardly biased death rates. All one needs is a sufficiently high number of deaths that cannot be matched to records of live persons. Furthermore, if mismatches occur because of faulty identifiers and/or the lack of relevant identifiers and this is more likely to occur among populations whose legal status is doubtful, then we could begin to suspect that the estimated advantage of foreign-born Mexicans and Other Hispanics is an illusion. For this to be the explanation of the Hispanic adult mortality advantage, though, the matching success rate of death and survey records would have to be at most 77% among foreign-born Mexicans and at most 54% among foreign-born Other Hispanics. These are remarkably poor success rates for a well-tested matching algorithm and are unlikely but not impossible, since with a small population base, it would take only a few hundred mismatches to produce such large downward biases. However, as we show later, what makes this explanation less plausible is that to account fully for the advantage among foreign-born Mexicans at least, the rate of mismatches should increase with age, a pattern that is inconsistent with the idea that illegal residence is at the root of faulty matching to begin with. And if not among foreign-born Mexicans, *those who are the most likely to count illegal migrants among their ranks*, why should the explanation based on mismatches hold for foreign-born Other Hispanics?

In sum, although we cannot rule out completely the hypothesis that the estimated Hispanic adult mortality advantage for foreign-born Mexicans and foreign-born Other Hispanics is due entirely to artifacts, we find all three sources of errors empirically unlikely and inconsistent with observable features of the mortality patterns.

In what follows, we explore and test the plausibility of Hypotheses 2, 3, and 4 to explain the Hispanic adult mortality advantage observed in the baseline models. We used a simplified representation of ethnic-nativity groups identical to the one in Model 2 of Table 1 and combined the male and female samples (see Palloni and Arias 2003). To account for gender differences in mortality levels, we included an additive term for gender. The strategy we used to choose among alternative representations of the data is based on the use of conventional p values and log-likelihood ratio statistics, combined with the Bayesian information criterion, or BIC (Raftery 1996). Only when we are reasonably certain about the performance (fit) of the model do we launch into an assessment of pertinent t statistics and related inferences. Information on estimates, standard errors, and associated log-likelihood ratio statistics appears in Tables 2 and 3 (for a detailed discussion of model-selection strategies, see Appendix D and Palloni and Arias 2003).

Assessing Hypothesis 2 (The Healthy-Migrant Effect)

The role of duration of residence in the United States. Through simulations, Palloni and Morenoff (2001) showed that health selection among migrants can have large, potent effects. Mechanisms through which those selection effects can take place were illustrated by Jasso et al. (2002). Thus, a sizable advantage may be completely attributable to initial differences in the health status of the populations being compared. Our baseline models suggest that the male and female advantages are far from trivial, since adult mortality rates among foreign-born Mexicans and Other Hispanics are 35%–47% lower than those among non-Hispanic whites. The corresponding relative risks translate into additional years of life expectancy at age 45 of approximately five to eight years of life. Palloni and Morenoff suggested that contrasts as large as and even larger than these can be created by healthy-migrant effects. But if this is so, it follows that healthy-migrant effects must leave at least some observable imprint in mortality data.

We argue next that the first imprint is a necessary convergence of the mortality of migrants and nonmigrants by the *duration of their stay in the United States*. The second imprint should be a sharp contrast between the mortality experiences of migrants residing in different areas of the United States.

Assimilation. Assume the existence of an advantage at the lowest duration of stay that can be produced by migrant health selection or by beneficial cultural endowments. Mortality rates for migrants and nonmigrants may become increasingly similar because the former group progressively adopts a more adverse profile of risk exposure—shaped by social, cultural, and behavioral factors—that resembles that of the nonmigrant population. Assimilation implies the jettisoning of favorable traits and the adoption of new ones in a trade-off with negative net health benefits for the migrant population in the following sense: had migrants preserved the original traits, their mortality levels would remain below those of the nonmigrant population, except at very old ages. In addition to assimilation, though, conditions associated with the migration experience per se, such as added stress and poor access to health care, contribute to less-favorable health and mortality profiles of the migrant population as their duration of stay increases.

Healthy-migrant effect. Suppose that the initial advantage of the migrant population is exclusively a function of health selection. On average, the migrant population is healthier than the reference domestic population and the population at origin. To simplify the argument, suppose also that the country of destination has a more favorable mortality regime and that migration does not lead to changes in migrants' risk profiles, that is, that there are no significant negative effects associated with assimilation or with the migration experience per se. What should be the patterns of contrast between the migrant and domestic population by duration of stay *when age effects are held constant*? The answer hinges on a subtle feature that is best illustrated with an example. Assume that there are two migrants of an identical age but whose duration of stay in the United States is different. The key difference between them is that they migrated at different ages. By assumption, the individual who migrated at a younger age has experienced the more-benign mortality regime of the country of destination longer and is therefore less selected for health-related traits than is the one who migrated at an older age. The consequence is that the effect of the duration of stay should be to attenuate the advantage experienced at higher durations; as in the case of assimilation, the advantage should be diluted as the duration of stay increases. Thus, the good news is that if healthy-migrant effects prevail, we should observe a decreasing advantage with duration of stay. The bad news is that the same pattern is compatible with an explanation that does not require health selection.

All this will be moot if there are no effects of duration of stay in the United States that follow the expected pattern. Table 2 offers information to help us make a judgment on this count.[11]

The first column of Table 2 displays the estimated effects associated with foreign-born Mexicans and Other Hispanics in our baseline model, that is, a model that includes only ethnicity and controls for socioeconomic traits. The second column of Table 2 presents estimates of the effects of dummy variables reflecting ethnicity (Mexican and Other Hispanic) and different durations of stay for migrants (less than 5 years, 5–9 years, 10–14

11. All models in Table 2 constrain the effects of Puerto Rican, Cuban, U.S.-born Mexican, and U.S.-born Other Hispanic to be identical. They all include controls for marital status, education, family income, and employment status (see Table 1). To avoid cluttering, this table displays only estimates that are associated with the ethnic-nativity groups of interest and with the variable or variables that are discussed in the text. All other estimates are omitted.

Table 2. Models Including Duration of Residence in the United States, State of Residence, and Age-at-Onset Effects (*SE* in parentheses)

Parameters	Model 1[a]	Model 2	Model 3	Model 4[b]	Model 5
Ethnicity-Nativity					
Foreign-born Mexican	–.340** (.055)	—	–.342** (.065)	–.222** (.071)	–.131 (.091)
Foreign-born Other Hispanic	–.690** (.080)	—	–.711** (.084)	–.585** (.105)	–.760** (.126)
Duration					
Foreign-born Mexican 1–4	—	–.332** (.066)	—	—	—
Foreign-born Mexican 5–9	—	–.465 (.253)	—	—	—
Foreign-born Mexican 10–14	—	–.414 (.233)	—	—	—
Foreign-born Mexican 15+	—	–.327** (.055)	—	—	—
Foreign-born Other Hispanic 1–4	—	–1.030* (.352)	—	—	—
Foreign-born Other Hispanic 5–9	—	–.507 (.365)	—	—	—
Foreign-born Other Hispanic 10–14	—	–.543** (.109)	—	—	—
Foreign-born Other Hispanic 15+	—	–.721** (.095)	—	—	—
Residence					
Non CA or TX	—	—	–.086* (.033)	—	—
Interaction × Foreign-born Mexican	—	—	–.545** (.133)	—	—
Age at onset					
Age ≥ 65	—	—	—	–.164* (.064)	–.165* (.061)
Interaction × Foreign-born Mexican	—	—	—	–.205** (.042)	–.372** (.061)
Interaction × Foreign-Born Other Hispanic	—	—	—	–.205** (.042)	.117 (.140)

(continued)

years, 15 or more years, and unknown) relative to the non-Hispanic white population. The results of this model are not entirely convincing. To begin with, the model itself does not fit well, at least relative to one that does not include the duration of stay in the United States. The chi-square statistic (2.0 with 6 degrees of freedom) is too small and statistically insignificant (see also the BIC statistic and associated information for Model 10 in Appendix D). Even if the estimated effects of a few duration dummy variables for ethnicity-nativity are marginally significant, the *pattern of effects* is inconsistent with either assimilation or health-selection effects. In fact, the pattern of effects is U-shaped,

(Table 2, continued)

Parameters	Model 1[a]	Model 2	Model 3	Model 4[b]	Model 5
Sample Size			39,013		
Events			3,253		
Log-Likelihood	–11,250.0	–11,249.0	–11,244.8	–11,245.2	–11,243.4
Diff. *df*[c]	—	6	2	2	1
Chi-square[d]	—	2.0	10.4	9.6	3.6
Upper Tail[e]	—	< .92	< .006	< .008	< .06

[a]All models were estimated on the pooled male and female sample. All models include dummies for ethnicity-nativity and controls for demographic and economic conditions, as defined in Table 1 (Model 2).

[b]In Model 4, the estimates of the effects for the two interaction effects with age at onset are constrained to be the same.

[c]Diff. *df* is the difference in the degrees of freedom between Model *j* and Model (*j* – 1) displayed in the preceding column.

[d]Chi-square is the chi-square statistic or –2 × (LLc – LLu), where LLc and LLu are the log-likelihood of the constrained and unconstrained models. Model 1 is the baseline against which Models 2–4 are contrasted. Model 5 is contrasted against Model 4.

[e]The upper-tail probability is the cumulated probability above the observed value of the chi-square statistic.

*p < .0114, or the absolute value of the t statistic is larger than 2 but smaller than 3.5; **p < .000116, or the absolute value of the t statistic is larger than 3.5.

with significant advantages at the shortest and longest durations of stay. The advantage at the longest duration of stay could reflect the attrition of unhealthier persons as the duration of stay increases and is certainly consistent with strong return-migration effects. Thus, this first test suggests that the data do not reveal the patterns one would expect to find from either selection effects or assimilation.

The role of region of residence. Migrant populations are not homogeneous, and neither is the degree of health selection within each of them. Differences in health selection across migrant populations should be reflected, in part, in the destination place: those who live in areas that demand largely unskilled labor and that offer greater and perhaps easier accessibility to points of entry (lowering the cost of migration) should be regions in which health selection is less rigorous.[12] The implication is that foreign-born Hispanics who live in or near the border areas should be less selected than should those who reside elsewhere. We tested the implication only with Mexicans. To do so, we created a variable for state of residence for the entire sample that attains a value of 0 if the individual lives in Texas or California and 1 otherwise.[13] To test the implication that the effects of being a foreign-born Mexican are larger for those who reside in states other than Texas and California, we created an interaction term using the dummy variable for state of residence and the dummy variable for foreign-born Mexicans. Estimates for the models that include the new variables are displayed in column 3 of Table 2.

We expected that the interaction term would be negative and significant, pointing to a higher advantage among foreign-born Mexicans residing in nonneighboring states. Our expectation was borne out with ample room to spare. Regardless of ethnic group, the effect

12. Although this inference is ad hoc, it can be retrieved from models of migration in which the risk of migration is made a function of both contrasts in the price of labor between the sending and destination regions and of the average costs of migration (see Jasso et al. 2002).

13. Because the bulk of illegal migration and entries from Mexico into the United States is concentrated in these two states, we thought we were justified in choosing these, instead of other states, as representative of areas containing the least-selected migrants. Because these are the most important "ports of entry," we reasoned that if the hypothesis fails to be rejected there, it would also fail more generally with other states.

of residing outside Texas or California is to reduce adult mortality by about (1 – exp(–.086)), or close to 8%. Although this contrast is of some interest by itself, it is the additional advantage for foreign-born Mexicans of residing in other states that concerns us here. This advantage is captured by the estimated coefficient of the interaction term, which is properly signed (negative), large, and statistically significant (–.545, t = ~4.1). The implications of this estimate are interesting. A foreign-born Mexican who lives in either Texas or California has a mortality rate that is exp(–.34), or about 71% as high as a non-Hispanic white who resides in either of the two states. But the contrast between foreign-born Mexicans and non-Hispanic whites who live in other states is much larger, namely, exp(–.34 – .55), or about a 41% lower mortality rate. Foreign-born Mexicans who live in states other than Texas and California have lower mortality rates than do foreign-born Mexicans who live in those two states not just by virtue of universal effects that apply to everyone who resides elsewhere (exp(–.086), about 8%), but because of an extra advantage that characterizes foreign-born Mexicans who live in other states (exp(–.55), about 58%).

Although the interaction effect is statistically significant with an exceedingly low significance margin (p < .0001) and is of a large magnitude, we think this test provides only weak support for the conjecture regarding selection effects. First, while the conventional goodness-of-fit statistic suggests that the addition of the two variables for state of residence improves the fit of the model (the last row of column 3 in Table 2), the BIC criterion associated with the model is somewhat unsatisfactory relative to a model that excludes those variables. (Compare the BIC values for Models 12 and 8 in Appendix D.) Second, the absolute magnitude and statistical significance of the foreign-born Mexican advantage is as strong as it was to begin with. Thus, although the test is suggestive, the results do not offer strong support for the idea that health-selection effects, reflected in the state of residence, account for the observed Hispanic advantage among foreign-born Mexicans and Other Hispanics: either there are no health-selection effects or, if they exist, they are not reflected well in state of residence, as assumed by the test.

Assessing Hypothesis 3 (Salmon-Bias Effect)

We now turn to an assessment of Type 1 and Type 2 return-migration effects. Because our data set did not enable us to distinguish empirically between these types, we refer to them as the “return-migrant effect” and treat them as a single bundle. The tests we discuss next involved contrasts of mortality rates and slopes across ethnic groups that should hold under both types of effects.

The importance of age effects. The first question is whether the estimates discussed before hold for the entire age span. If the model estimated earlier is true, then the answer to this question is obviously affirmative. But because we do not know for sure, we should estimate an alternative model allowing the effect of being (foreign-born) Mexican and Other Hispanic to be different at ages above and below some threshold, say age 65. The justification for this argument is as follows: if return-migration effects are influential, we would expect the advantage to be proportionately larger at older ages. Furthermore, because the magnitude of these effects is a function of return-migration rates, the advantage is more likely to occur among foreign-born Mexicans than among foreign-born Other Hispanics whose countries of origin are less easily reachable. Thus, we expect the age differences in the advantage to be trivial for Other Hispanics but significant for Mexicans.[14]

14. Behind this reasoning may lie an important simplification. Journalistic reports have suggested that return migration for some non-Mexican Hispanic groups, such as Salvadorans and Guatemalans, was fairly high from 1995 to 2000. The flows were fueled, in part, by massive deportations, but at least some of them originated in voluntary repatriation, so it should follow that age differences could also be important among Other Hispanics.

To test this conjecture, we defined a new dummy variable as 0 if the age at onset of the study was younger than 65 and set it equal to 1 if the age at onset was 65 and older. We then estimated two models, one in which the effect of an interaction term between the variables for ethnicity-nativity (foreign-born Mexican and Other Hispanic) and the dummy variable for age group are identical for both Mexicans and Other Hispanics and one in which the effects of the interaction term are unconstrained. The results are displayed in the last two columns of Table 2. The constrained model (column 4) yields a negative and significant effect of the interaction term between the dummy variable for older age and the dummy variables for foreign-born Mexican and Other Hispanic. This finding means that, as expected, if there is a return-migrant effect, the advantage is larger for those who were aged 65 and older at the beginning of the study. But the main effects for foreign-born Mexican and Other Hispanic still persist. It is in the unconstrained model (the last column of Table 2) where we found three features expected by our conjecture: (1) the effect of older age applies to foreign-born Mexicans only, (2) the advantage among foreign-born Mexicans vanishes and is replaced by pure age effects, and (3) the advantage enjoyed by foreign-born Other Hispanics remains unchanged. These three regularities are exactly as one would conjecture if there are significant return-migrant effects and if these effects are more likely among Mexicans than among Other Hispanics.

Can these effects be interpretable as data artifacts? The only source of data artifact that could plausibly be brought to bear is the overstatement of ages. Because all models were estimated constraining age and slope effects and in all cases the constraint holds, there is no reason to suspect large effects of the overstatement of ages.

Yet, despite its apparent success, the model has three drawbacks. First, although the tests of significance suggest the existence of age effects, the goodness of fit of the preferred model is only marginally better, if at all, than that of the constrained model. Thus, the estimate of the interaction effect with foreign-born Mexican (–.372) is large, negative, and statistically significant, and the difference between it and the interaction term for foreign-born Other Hispanic is of the opposite sign, large (.489), and also statistically significant (t = ~3.5). But the likelihood-ratio test for the contrast between Models 4 and 5 in Table 2 is only 3.6. With 1 degree of freedom, this is only marginally significant (the last row of last column, Table 2).

Second, the results of the test reflect a shift in mortality levels for older cohorts. Although this is an expected consequence of return-migration effects, it can also be the outcome of cohort changes in mortality. Although the latter possibility is unlikely because the shift would require a deterioration of mortality levels at older ages for those who were members of younger cohorts, it is an unsettling possibility. The key is that the interpretation of the observed pattern may not be unique and Model 5 could be a good representation of empirical relations without implying the existence of any return-migration effects.

The third and final problem is that the age we chose as a cut point (65) is arbitrary, and there is no reason to select it over many other potential candidates that may yield completely different results.

Mortality slopes. There is still an unexplored possibility. Rather than result in a shift of rates after a certain age, return-migration effects should exert a gradual influence, spread out over a large range of duration; this possibility should be reflected in a reduced slope of the mortality pattern. The gradual depletion of individuals in bad health leads, via return migration of the unhealthy spread out over the duration of the study, to a mortality pattern with a downwardly biased slope.[15] Our argument is that we need to make room for effects operating not just on the level of mortality but also on the rate of change

15. This would be the case if and only if Type 2 effects were proportionally greater at older ages, as they should be under conditions of unrestricted return migration.

of mortality risks with duration in the study. If this is so, Gompertz's slope must vary across ethnic groups, even after constraining age effects to be identical to what is common in the slope effect across ethnic groups. More formally, in Eq. (1a), we replaced γ by $\gamma = \gamma_o + \zeta Z_e$, where γ_o is a constant, Z_e is a dummy variable reflecting membership in ethnic group e, and ζ is an effect on the slope. We then constrained δ, the effect of age (rescaled) at the beginning of the follow-up period, to be equal to γ_o. Thus, we let the slope change but not the constant slope and the age effect. The slopes should be the lowest in groups that are the most affected by return migration. Unlike the age-related shift in the advantage documented earlier, a slope effect is more difficult to account for by discrete cohort effects; it does not require an arbitrary choice of age as a cut point; and, what is not trivial, it cannot be confused with the impact of the overstatement of ages, *since it reflects only duration effects*. Letting the slope be a function of covariates is tantamount to saying that the underlying hazard model is not proportional and that the effects of covariates are not captured by constant shifts of the log of hazards.

To test this possibility, we estimated two models in which the slope of the Gompertz curve is a function of ethnicity. The first model seeks to determine whether the slope is lower for foreign-born Mexicans than it is for everybody else, as it should be if a return-migration effect is stronger among Mexican migrants than among all other groups. The second model tests the same implication but for both foreign-born Mexicans and foreign-born Other Hispanics.

The first two columns of Table 3 contain the results for these two models. The estimated effect of being Mexican on the slope is, as expected, negative and statistically significant ($t = {\sim}3.0$). All the information for this model indicates that it fits the data well, marginally better than a model in which the slope is constrained to be a constant (for example, compare with Model 8 in Appendix D). Finally, and more important, the effects of being foreign-born Mexican vanish, while those associated with being Other Hispanic remain strong.

If this pattern were also evident for Other Hispanics, our explanation would lose credibility, since such a pattern is expected only if the observed mortality advantage is accounted for by return migration. To the extent that return migration among Other Hispanics is much less likely, we should not expect an effect on the slope similar to that just estimated among foreign-born Mexicans. In fact, this is the case. The third column of Table 3 displays estimates of a model in which we allowed the effects of being either foreign-born Mexican or foreign-born Other Hispanic on the slopes of the mortality curve, but only the former are in the expected direction and significant. The effect of being Other Hispanic on the slope is positive and insignificant. Furthermore, the estimated advantage increases; it does not become weaker. Finally, the fit of the model (see Appendix D) is poor when compared with its simpler version. These are all signs that return-migration effects cannot possibly explain the mortality advantage among foreign-born Other Hispanics.

These findings lead to the following two propositions. First, the advantage of foreign-born Mexicans is largely accounted for by the smaller slope of their mortality curve, a telltale sign of return-migrant effects. Neither duration of residence nor residential patterns account for much of the observed advantage, suggesting that health-selection effects may be less important than return-migration effects. Second, the advantage for foreign-born Other Hispanics is remarkably robust. It is not related to return migration (as reflected in the slope or age effects), and it does not weaken when the effects of duration of residence are controlled.

Assessing Hypothesis 4 (Cultural Effect)

As we stated at the beginning of this article, the data set available to us contained no information on individual connections and social networks, integration into a community,

Table 3. Models Including Duration Tests for Slopes and Cultural Effects (*SE* in parentheses)

Parameters	Model 1[a]	Model 2[b]	Model3[b]	Model 4[c]
Baseline				
Constant	–6.219 (.042)	–6.250 (.045)	–6.234 (.049)	–6.194 (.044)
Slope				
Constant	.070 (.0013)	.071 (.0011)	.070 (.0014)	.070 (.0013)
Foreign-born Mexican	—	–.0070** (.0024)	–.0067** (.0025)	—
Foreign-born Other Hispanic	—	—	.0072 (.0047)	—
Ethnicity/Nativity				
Foreign-born Mexican	–.340** (.055)	–.089 (.136)	–.102 (.132)	–.358** (.078)
Foreign-born Other Hispanic	–.690** (.080)	–.686** (.082)	–.943** (.183)	–.700** (.087)
Isolation Index				
First quartile	—	—	—	—
Second quartile	—	—	—	–.078** (.017)
Third quartile	—	—	—	–.024 (.036)
Fourth quartile	—	—	—	.004 (.045)
Sample Size	39,013			
Events			3,253	
Log-Likelihood	–11,250	–11,249	–11,245	–11,248
Diff. *df*[d]	—	—	—	3
Chi-square[e]	—	—	—	5.4
Upper Tail[f]	—	—	—	< .15

[a]All models were estimated on the pooled male and female sample. All models include dummy variables for ethnicity-nativity and controls for demographic and economic conditions as defined in Table 1 (Model 2)

[b]Model 1, our baseline model, is not nested in Models 2 or 3, and conventional chi-square statistics do not apply.

[c]Model 1 is nested in Model 4 and Model 3, and conventional tests apply.

[d]Diff. *df* is the difference in the degrees of freedom between model j and model $(j – 1)$ displayed in the preceding column.

[e]Chi-square is the chi-square statistic or $–2 \times$ (Llc – Llu), where LLc and LLu are the log-likelihood of the constrained and unconstrained models.

[f]The upper-tail probability is the cumulated probability above the observed value of the chi-square statistic.

*$p < .0114$, or the absolute value of the t statistic is larger than 2 but smaller than 3.5; **$p < .000116$, or the absolute value of the t statistic is larger than 3.5.

or the like. Thus, it was impossible to design a rigorous test of the cultural hypothesis, according to which some of the Hispanic advantage that remains to be explained is associated with cultural conditions. To bridge the gap, we conducted two decidedly humbler tests. The first relies on the argument that the primary line of social protection is derived from an individual's family ties. If so, the universally protective effects of marriage ought to be even stronger. This argument suggests that the health benefits among married migrants must exceed those of the general population. The model we estimated (results not shown) included an interaction term for marital status and for being either foreign-born Mexican or foreign-born Other Hispanic. But although the effects are negative as expected (the protective effects of marriage are higher among these groups), they are statistically insignificant, and the model's fit is mediocre relative to other models.

The second test sought to control for the type of community in which migrants live. If the advantage is associated with community ties, communities that are ill prepared to offer an abundant supply of ties will not be good for migrants. Communities that do not score high on an index of segregation or isolation of migrant groups are not fertile grounds in which to cultivate social ties and to establish social networks. Furthermore, communities that lack such networks may have a negative impact by magnifying immigrants' sense of isolation, loneliness, or discrimination. If the hypothesis of cultural and social effects has any relevance, the mortality advantage should vanish once the type of community in which migrants live is controlled.

Although we just showed that the advantage associated with being foreign-born Mexican disappears when we account for slope effects, we estimated a model that included indicators of isolation with no slope effects to assess whether it can account for observed patterns, as does the one with variable slopes. The last column of Table 3 shows pertinent estimates of the models. Interaction terms between foreign-born Mexican and Other Hispanic and the index of isolation of the community within which each individual lives were also estimated (results not shown; see Appendix C for a definition of the variables). Once again, although the effects are in the expected direction, neither is large enough to attain more than modest levels of statistical significance. And, what is worse, neither the advantage associated with Mexicans nor the advantage associated with Other Hispanics changes much. Indeed, they are as strong as they were in our baseline models.

Modest as they may be, neither of these tests supports the validity of the cultural explanation. Surely, our failure to detect an influence of isolation on the estimated Hispanic adult mortality advantage should not be interpreted to mean that social and cultural factors are immaterial. They are, indeed, important because they modify the risks of mortality. It is just that they do not help us account for the difference in mortality risks among Hispanics.[16]

Forcing a Choice of Model

We believe that a model with a variable slope is the most appropriate for the data. It accounts for the advantage of foreign-born Mexicans, although without simultaneously reducing the advantage of foreign-born Other Hispanics. It is an appropriate formulation that captures the presence of return-migration effects and is much less amenable to be interpreted as a result of data artifacts. Tests for unmeasured frailty suggest that an admittedly narrow class of frailty effects did not change our results at all, thus increasing our confidence that what we observed is less likely to be produced by unobserved factors. A more serious threat to the interpretation we offer is associated with the possibility that a similar pattern of results would be observable because of the overstatement of ages (Dechter and Preston 1991; Preston et al. 2003; Rosenwaike 1991; Rosenwaike and Preston 1983). But this possibility is unlikely for three reasons. First, the declaration of

16. In addition to the model defined earlier, we tested a number of specifications that account for parametric forms of unmeasured heterogeneity, but in no cases did we find any evidence that they changed the estimates of key parameters or improved the fit of the models (see Palloni and Arias 2003).

Table 4. Comparison of Expected and Observed Death Counts From the Competing Models

Ethnic Group	Exposed Individuals	Observed Deaths	Expected Deaths Model 1	Model 2	Model 3	Model 4
Non-Hispanic Whites	22,554	2,250	2,218	2,214	2,218	2,216
Foreign-born Mexicans	3,706	196	190	192	190	192
Foreign-born Other Hispanics	3,022	97	94	97	94	94
Remaining Hispanics	9,745	701	693	693	692	693

Notes: Model 1 refers to the baseline or Model 8 in Appendix D (ethnicity and controls for socioeconomic status). Model 2 is Model 1 with added variables for the dummy variable for age at onset of the study (> 65). Model 3 is Model 1 with added variables for state of residence. Model 4 is Model 1 with the slope a function of one covariate (foreign-born Mexican).

ages among the living population is better in NHIS than in censuses and, by the very nature of the data set, the overstatement of ages of the deceased is less of a problem. Second, the tendency to overstate ages has been attributed to Hispanics in general, not to particular groups. It is hard, then, to reconcile this interpretation with the observed pattern in which negative slope effects are prevalent only among the non-U.S.-born Mexican population but not among Other Hispanics. Third and more important, the overstatement of ages should not lead to slope effects at all. The effects of age overstatement should be reflected in estimates of age effects that are lower than estimates of the slope.

Although the variable-slope model is theoretically sound, is resistant to artifacts, and appears to be statistically preferable to others, how much better is it than its competitors? So far, we have relied on conventional *p* values associated with chi-square statistics and on the less-liberal BIC values. Neither of these values provides a concrete metric of fit, one with a straightforward interpretation and with practical implications. One solution is to compare the predicted number of deaths during the period of observation derived from the models we consider as top choices. Table 4 provides the observed and expected counts of deaths for each of four ethnic groups computed from each of four competing models. A cursory examination of the tables reveals the key problem: they are all so good at predicting the count of deaths by ethnic groups that it is impossible to tell them apart.

Thus, our main conclusion must be necessarily subdued: although we have a theoretical preference for the more elegant model with a variable slope, it does not perform significantly better than competing models in accounting for observed death counts. We clearly need additional data to justify our preference.

Reconsidering the Salmon Bias: Comparing Migrants

One of the main conclusions of this article is that the bulk of the foreign-born Mexican adult mortality advantage is related to the return migration of those who are in poor health. We reached this conclusion using indirect, rather than direct, evidence. To reinforce it, we now briefly describe the results of an evaluation of data that can bring us close to a direct type of evidence.

The ideal test for the return-migrant hypothesis is to compare the mortality of recent return migrants to the mortality of migrants who remained in the country of destination. Such a comparison is difficult, since there is no follow-up of return migrants. However, we do have information on the heath status of adults aged 50 and older and their surviving spouses, who were interviewed in Mexico during 2000. This data set—the Mexican Health and Aging Study (MHAS; see Soldo, Wong, and Palloni 2002)—also provides information on the migrant status of individuals who reside in Mexico and a limited migration history. With this data set, we could compare the self-reported health status of return migrants with the self-reported status of individuals in the NHIS-MCD sample at various points during the follow-up. If return-migration effects are strong, the self-reported health status of return migrants should be worse than that of migrants who stayed in the United States. For this comparison, we selected the subsample of adults aged 50 and older who were interviewed by MHAS, who resided in the United States, and who had returned to Mexico in the 10 years before the survey.[17] We compared this group to three NHIS-MCD samples. The most inclusive among them (Sample 1) consisted of all foreign-born Mexicans in the NHIS-MCD baseline study. The second sample (Sample 2) consisted of all foreign-born Mexicans who were in the initial NHIS-MCD sample and who were not matched to death records for deaths that occurred before the midpoint of the follow-up period. The third sample (Sample 3) consisted of all those who were enumerated in the baseline survey and who were not matched to death certificates during the entire follow-up period. Because of the progressive removal of those who died during the follow-up, the health-status composition of Sample 1 should be the worst and that of Sample 3 should be the best. Although all three samples included return migrants, return migrants were the most heavily represented in Sample 3 because all individuals who died during the follow-up period were removed. Neither of these three samples corresponds to a true sample of stayers, but Sample 1 should provide a lower bound for the health status of stayers, whereas Sample 3 should provide an upper bound.

The comparison should be suggestive and not decisive for a number of reasons. First, we do not know much about the biases inherent in self-reports or about how they change with ethnicity, with place of residence, and with the duration of stay in the United States. Second, the samples we compared are not consistent in age or in the timing of migration. Third, the NHIS-based self-reports included individuals who, in due course, became return migrants, and it is thus not a sample of true stayers. Fourth, return migrants who were in the worst health are probably not represented among the surviving return migrants in Mexico because they may have died before they could become part of the sampling frame of MHAS. Most of these difficulties, however, play in our favor, since the differences in self-reported health status between the MHAS sample and the NHIS-MCD sample will underestimate the differences between a true sample of recent returnees and a true sample of stayers.

Table 5 displays the percentages of individuals who self-reported as being in bad health and in fair or bad health in three age groups in all four samples that were previously considered. As expected, the health distribution of Sample 1 is the worst and that of Sample 3 is the best. This finding simply reflects the fact that self-reported health is a moderately good predictor of mortality. Also as expected, the health status of those at older ages deteriorated in all the samples.

The comparison with the MHAS sample is a bit difficult because the latter is based on a small number of cases. However, as would be expected if return migrants were selected among those in bad health, the MHAS respondents' health status was worse than the health statuses of the respondents in any of the NHIS-MCD samples. Although consistent with the return-migration conjecture, this finding is fragile for two reasons. First, the sample frequencies are too small to justify more than modest enthusiasm. Second, a comparison between the MHAS sample and the subsample of return migrants shows that the health status of the latter was marginally better than the health status of the entire MHAS sample. While this pattern is not inconsistent with return-migration

17. It is obviously better to choose a subsample of more recent returnees. Although it is feasible to do so, the frequencies involved are too small to make meaningful comparisons. Choosing an interval of 10 or fewer years was the best compromise we could find to resolve the tension between the sample size and the recency of return migration.

Table 5. Percentage Who Self-Reported as Being in Poor (P) and in Poor and Fair Health (P + F) for Three Different NHIS-MCD Samples and a Subsample of U.S. Return Migrants Living in Mexico

	NHIS-MCD Samples[a]								
	Sample 1		Sample 2		Sample 3		MHAS Subsample		
Age Group	P	P + F	P	P + F	P	P + F	N^b	P	P + F
50–59	8.5	28.0	7.7	27.4	7.6	26.2	112	11.6	50.8
60–74	11.2	35.4	10.1	33.2	9.7	32.1	63	15.9	63.4
75 and Older	16.2	40.7	11.8	35.0	11.4	34.5	10	20.1	60.1

[a]See the text for definitions of Samples 1, 2, and 3 and for the definition of the MHAS subsample.

[b]Refers to the number of cases (unweighted) in each age group.

effects—indeed, they could be found when initial health selection in the immigration flow to the United States is combined with return-migration effects—it begs for an explanation. Until this explanation is offered, the evidence just produced must be taken as circumstantial. Matters may be more complicated than what we assumed in this simple test.

SUMMARY AND CONCLUSIONS

After a fairly dense battery of tests, we can draw the following conclusions:

1. The Hispanic adult mortality advantage is not "Hispanic." Rather, it is a feature only of foreign-born Other Hispanics and foreign-born Mexicans—not of Puerto Ricans or Cubans, whether born in the United States or abroad.

2. The foreign-born Mexican and Other Hispanic adult mortality advantage is not trivial. It amounts to experiencing mortality rates that are 35% to 47% lower than those experienced by non-Hispanic whites. In turn, these differences translate into approximately five to eight years of additional life expectancy at age 45.

3. The behavior of mortality slopes produces strong signs of return-migration effects for foreign-born Mexicans but not for Other Hispanics. Although the model we used to confirm this pattern fit the data as well as or only marginally better than did competing models, our conclusion is robust to a class of unmeasured heterogeneity and received additional support from a comparison of several NHIS-MCD samples and the MHAS sample. Indications of the presence of health-selection effects are reduced and circumscribed to effects of state of residence.

4. The observed advantage favoring Other Hispanics persists even after indirect consequences of healthy-migrant effects (duration of stay, state of residence) are accounted for and is resistant to age and slope effects, as well as to unmeasured heterogeneity. This does not mean that healthy-migrant effects are absent, but that, if they exist, they are not reflected strongly enough in the mechanisms we were able to identify (duration or residence effects).

5. The cultural hypotheses received no support. We uncovered effects suggesting that those who live in ethnically more cohesive communities have lower mortality, as one would expect from the cultural hypotheses. But these conditions do not account for the Hispanic advantage, nor do they alter the effects of membership in a group. It is not because foreign-born Mexicans or Other Hispanics have a higher propensity to live in cohesive communities that they experience lower mortality than do non-Hispanic whites. And

it is not because there are extra gains accruing from residence in those communities among some Hispanics that there are mortality advantages.

This has been a partially successful exploration of the problem. We were able to justify a model that accounts for part of the advantage and attribute it to return-migration effects. However, the preferred model that spawned this interpretation neither rests on robust, uncontested grounds nor is complete, since part of the advantage—the part associated with Other Hispanics—remains thoroughly unexplained.

Appendix A. Description of the Variables Used in the Analyses

Ethnicity: non-Hispanic white, Cuban, Mexican, Puerto Rican, and Other Hispanics. Ethnicity was self-identified at the time of the interview by persons aged 18 and older.

Age: age stated in years at the time of the interview.

Delta: age measured in the number of years since the lower bound for which the hazard is applicable. For example, for the sample of those aged 35 and older, delta = age – 35.

Marital status: married, not married, and unknown. Not married includes never-married, separated, divorced, and widowed persons. This variable refers to the marital status of the individual at the baseline.

Education: less than high school, high school graduate, more than high school, and unknown. This variable refers to the educational attainment of the individual at the baseline.

Family income: four empirical quartiles of the family income distribution, from the first quartile (the lowest 25th percentile of the income distribution) to the fourth quartile (the highest 25th percentile of the family income distribution). This variable refers to family income reported at the time of the interview.

Employment status: currently employed, currently unemployed, and not in the labor force. This variable refers to the employment status of the individual at the time of the interview.

Nativity/duration of residence: foreign born in the United States for less than 5 years, foreign born in the United States for 5–9 years, foreign born in the United States for 10–14 years, foreign born in the United States for 15 or more years, U.S. born, and unknown. Finer categories of duration of residence were not possible because these are the categories reported in the NHIS. This variable refers to the nativity/duration status of the individual at the time of the interview.

Isolation index was generated from the Census STF1 1990 file and appended to the NHIS-MCD data set by matching the data sets by FIPS state and county codes. It is a physical segregation or measure of exposure. This measure can be interpreted as the probability that a selected minority member is exposed to only other members of his or her minority group. It is estimated as follows: ${}_xP_x = \Sigma_i\,[x_i / X] \times [x_i / t_i]$, where x_i and t_i are the numbers of minority members and the total population in unit i, respectively. The unit of analyses in this case is the county tract. X represents the total number of x minority members in the county. The index ranges from 0 to 1 and may be interpreted as the probability that a randomly selected member of a particular ethnic group lives in the same geographic area as a coethnic (Massey and Denton 1988).

For the purposes of this study, X refers to the total Hispanic population in unit i (county tract), and the index is summed to the county level. The distribution of the index is broken down into four empirical quartiles, where the first quartile refers to the lowest 25th percentile of the index's distribution (or the least amount of segregation) and the fourth quartile to the highest 25th percentile of the index's distribution (or the greatest segregation).

State of residence: a two-category variable reflecting residence in California or Texas (0) and residence in any other states (1). It was constructed to refer to the state residential status of Mexican Americans at the time of the interview. California and Texas are border states with Mexico.

Appendix B. Sample Size and Mortality Outcome Characteristics, Men and Women, by Age Group

	Non-Hispanic White		Cuban		Mexican		Puerto Rican		Other Hispanic	
Age Group	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Age 35 and Older										
Sample size	10,389	12,176	664	800	4,047	4,463	860	1,123	1,990	2,792
Alive	9,269	11,045	588	724	3,768	4,243	795	1,059	1,879	2,668
Deaths	1,120	1,131	76	76	279	220	65	64	111	124
Age 35–59										
Sample size	6,503	7,007	394	447	3,150	3,295	661	841	1,527	2,051
Alive	6,324	6,886	382	439	3,067	3,238	639	815	1,503	2,030
Deaths	179	121	12	8	83	57	22	26	24	21
Age 60 and Older										
Sample size	3,886	5,169	270	353	897	1,168	199	282	463	282
Alive	2,989	4,187	208	290	727	1,019	160	247	378	645
Deaths	897	982	62	63	170	149	39	35	85	96

Appendix C. Descriptive Statistics of the Sample of Men and Women Aged 35 and Older

	Non-Hispanic White		Cuban		Mexican		Puerto Rican		Other Hispanic	
Age Group	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Age										
35–59	63.0	57.5	57.8	53.3	78.2	74.3	76.9	74.5	76.5	73.5
60 and older	37.0	42.6	42.2	46.7	21.8	25.7	23.1	25.5	23.5	26.5
Marital Status										
Married	81.8	65.0	78.8	56.6	83.5	65.8	77.6	51.8	80.9	58.7
Not married	18.1	34.8	21.1	43.5	16.4	34.0	22.3	48.0	19.0	41.1
Unknown	0.1	0.2	0.1	0.0	0.1	0.2	0.1	0.2	0.1	0.3
Education										
< High school	20.5	22.0	38.5	42.6	55.8	59.1	47.9	51.7	30.8	38.0
High school	33.7	41.2	23.9	29.1	22.0	24.7	28.5	26.8	27.6	30.0
> High school	45.3	36.2	37.1	27.5	20.6	15.2	22.4	20.6	40.9	31.3
Unknown	0.5	0.6	0.5	0.9	1.6	1.0	1.3	0.9	0.7	0.7
Family Income Distribution										
First quartile	13.9	22.1	23.8	29.7	28.3	36.7	30.0	42.3	18.9	29.3
Second quartile	23.5	25.3	32.6	31.6	30.0	27.0	26.2	24.2	26.4	24.7
Third quartile	29.0	27.0	24.0	20.3	24.3	22.1	23.1	18.4	27.3	24.4
Fourth quartile	33.6	25.6	19.7	18.5	17.4	14.3	20.7	15.1	27.5	21.7

(continued)

(Appendix C, continued)

Age Group	Non-Hispanic White Male	Non-Hispanic White Female	Cuban Male	Cuban Female	Mexican Male	Mexican Female	Puerto Rican Male	Puerto Rican Female	Other Hispanic Male	Other Hispanic Female
Employment Status										
Employed	66.4	47.0	63.2	43.0	70.7	46.2	62.8	38.2	73.2	53.0
Unemployed	2.0	1.6	3.1	2.2	4.8	3.2	2.4	2.3	3.5	2.9
Not in the labor force	31.6	51.4	33.4	54.8	24.5	50.7	34.8	59.5	23.4	44.0
Nativity/Duration of Residence in the United States										
< 5 years	0.4	0.3	5.1	4.9	3.1	2.9	4.5	5.8	6.2	6.7
5–9 years	0.2	0.2	6.2	4.2	4.6	3.2	4.7	4.3	11.4	9.4
10–14 years	0.4	0.3	14.7	8.2	6.6	5.7	4.4	5.3	10.0	9.6
15 or more years	3.5	4.6	59.6	67.5	31.1	29.5	61.9	61.0	34.9	38.1
U.S. born	95.4	94.6	13.7	13.5	53.3	57.4	21.5	20.1	36.2	34.8
Unknown	0.0	0.1	0.8	1.8	1.3	1.4	2.9	3.5	1.4	1.4
Isolation Index Distribution (0–1)										
First quartile	39.6	40.3	6.0	5.0	2.1	1.8	4.2	5.3	7.5	7.0
Second quartile	33.1	33.3	9.6	11.3	10.6	9.7	20.3	16.5	20.3	21.9
Third quartile	19.2	19.5	14.6	11.1	32.9	36.1	42.3	41.3	34.4	34.8
Fourth quartile	8.2	6.9	69.8	72.6	54.4	52.4	33.1	37.0	37.8	36.3

Appendix D. Quantities for Calculating BIC Statistics in All Models Estimated With Combined Samples

Model J	Table Location	Remarks About Model J	Kj	Pj	–LLj	LRTj	–BICj
0	Not displayed	Null model (one constant)	1	—	13,651	—	—
1	Not displayed	Two baseline (Gompertz) parameters	2	1	11,447	4,408	4,400
2	Not displayed	As Model 1 plus gender	3	2	11,377	4,548	4,531
3	Not displayed	As Model 2 + four parameters for ethnicity	7	6	11,364	4,574	4,525
4	Not displayed	As Model 3 + one parameter for nativity	8	7	11,347	4,608	4,551
5	Not displayed	As Model 2 + eight parameters for ethnicity/nativity	11	10	11,344	4,614	4,533
6	Not displayed	As Model 5 + nine parameters for SES controls	20	19	11,248	5,806	4,652
7	Not displayed	As Model 2 + three parameters for ethnicity/nativity	6	5	11,350	4,602	4,561
8	Table 2	As Model 7 + nine parameters for SES controls	15	14	11,250	4,802	4,689

(continued)

(Appendix D, continued)

Model J	Table Location	Remarks About Model J	K*j*	P*j*	–LL*j*	LRT*j*	–BIC*j*
9	Not displayed	As Model 8 + four parameters for gender interactions with foreign-born Mexicans and Other Hispanics	19	18	11,346	4,610	4,464
10	Table 2	As Model 8 + six parameters for duration in the United States	21	20	11,249	4,804	4,642
11	Table 2	As Model 8 + two parameters for residence and interaction with foreign-born Mexicans	17	16	11,245	4,812	4,683
12	Table 2	As Model 8 + two parameters for age ≥ 65 at onset and one interaction for foreign-born Mexicans and Other Hispanics	17	16	11,245	4,812	4,683
13	Table 2	As Model 12 but with two interaction terms	18	17	11,244	4,814	4,677
14	Table 3	As Model 8 + one extra parameter for the slope (effect of foreign-born Mexicans)	16	15	11,249	4,804	4,683
15	Table 3	As Model 8 + two extra parameters for the slope (effect of foreign-born Mexicans and Other Hispanics)	17	16	11,248	4,806	4,677
16	Table 3	As Model 8 + three parameters for isolation of community of residence	18	17	11,248	4,806	4,669
17	Not displayed	As Model 16 + three interaction terms for isolation and foreign-born Mexicans and Other Hispanics	21	20	11,246	4,810	4,648

Notes: This table summarizes all the necessary quantities for the calculation of log-likelihood ratio statistics and BIC statistics for all relevant models. The estimates for a few of the models included here are not shown in the other tables because they are less important substantively. All likelihood-ratio statistics and corresponding BIC values for the models are calculated with reference to the null model, one in which mortality is represented by only one parameter. Furthermore, a comparison of BIC values for Models 5 and 6 suggests a representation that includes separate effects for only three Hispanic groups (foreign-born Mexicans, foreign-born Other Hispanics, and all other Hispanics). Also, our choice of a pooled male and female sample is further justified by a comparison of Models 8 and 9. Model 8 includes only the additive effects of gender, whereas Model 9 also includes interaction terms between gender and two variables for foreign-born Mexican and Other Hispanic. If there were important gender differences in relations implicating ethnicity, Model 9 should fit better. As expected, it does not. Models 5–9 correspond to the set of models out of which the main baseline is selected (Model 8). Models with lower (more negative) BIC values should be preferred. Kj = the number of free parameters in Model j; Pj = the number of degrees of freedom associated with the test comparison of Model j and the null model; –LLj = the (–1) × log-likelihood of Model j; LRTj = the log-likelihood ratio test statistic for the contrast between Model j and the null model, or –2 × (LL0 – Llj); and BICj = the quantity (–LRTj + Pj × lnE), where E is the number of events. Sample size *N* = 39,014. Number of Events = 3,253.

REFERENCES

Abraido-Lanza, A.F., B.P. Dohrenwend, D.S. Ng-Mak, and J.B. Turner. 1999. “The Latino Mortality Paradox: A Test of the ‘Salmon Bias’ and Healthy Migrant Hypotheses.” *American Journal of Public Health* 89:1543–48.

Adler, N.E. and J.M. Ostrove. 1999. “Socioeconomic Status and Health: What We Know and What We Don’t.” Pp. 3–16 in *Socioeconomic Status and Health in Industrial Nations: Social, Psychological and Biological Pathways,* edited by N.E. Adler, M. Marmot, B.S. McEwen, and J. Stewart. *Annals of the New York Academy of Sciences,* Vol. 896.

Arias, E. 1998. “The Demography of Assimilation: The Case of Cubans in the United States.” Unpublished doctoral dissertation. Department of Sociology, University of Wisconsin–Madison.

———. 2002. “United States Life Tables, 2000.” *National Vital Statistics Reports* 51(3). Hyattsville, MD: National Center for Health Statistics.

Becker, T.M., C. Wiggins, C.R. Key, and J.M. Samet. 1988. “Ischemic Heart Disease Mortality in Hispanics, American Indians, and Non-Hispanic Whites in New Mexico, 1958–1982.” *Circulation* 78:302–309.

Dechter, A. and S.H. Preston. 1991. “Age Misreporting and its Effects on Adult Mortality Estimates in Latin America.” *Population Bulletin of the United Nations* 31–32:1–16.

Hummer, R.A., R.G. Rogers, S.H. Amir, D. Forbes, and W.P. Frisbie. 2000. “Adult Mortality Differentials Among Hispanic Subgroups and Non-Hispanic Whites.” *Social Science Quarterly* 81:459–76.

Hummer, R.A., R.G. Rogers, C.B. Nam, and F.B. LeClere. 1999. “Race/Ethnicity, Nativity and U.S. Adult Mortality.” *Social Science Quarterly* 80:136–53.

Jasso, G., D. Massey, M. Rosenzweig, and J. Smith. 2002. “Immigrant Health: Selectivity and Acculturation.” Discussion paper for the Conference on Racial and Ethnic Disparities in Health, National Academy of Science.

LeClere, F.B., R.G. Rogers, and K.D. Peters. 1997. “Ethnicity and Mortality in the United States: Individual and Community Correlates.” *Social Forces* 76:169–98.

Liao, Y.L., R.S. Cooper, G. Cao, R. Durazo-Arvizu, J.S. Kaufman, A. Luke, and D.L. McGee. 1998. “Mortality Patterns Among Adult Hispanics: Findings From the NHIS 1986 to 1990.” *American Journal of Public Health* 88:227–32.

Markides, K.S. 1983. “Mortality Among Minority Populations: A Review of Recent Patterns and Trends.” *Public Health Reports* 98:252–60.

Markides, K.S. and J. Coreil. 1986. “The Health of Hispanics in the Southwestern United States: an Epidemiologic Paradox.” *Public Health Reports* 101(3):253–65.

Massey, D.S. and N.A. Denton. 1988. “The Dimensions of Residential Segregation.” *Social Forces* 67:281–315.

Mendes de Leon, C. and T. Glass. 2002. “The Role of Social Networks and Personal; Resources in Ethnic Disparities in Aging Health.” Discussion paper for the Conference on Racial and Ethnic Disparities in Health, National Academy of Science.

National Center for Health Statistics (NCHS). 1989. “Design and Estimation for the National Health Interview Survey, 1985–94.” *Vital and Health Statistics, Data Evaluation and Methods Research*, Series 2 (110). Hyattsville, MD: National Center for Health Statistics.

———. 2000. “National Health Interview Survey Multiple Cause of Death Files—2000 Linkage.” Hyattsville, MD: National Center for Health Statistics.

Novello, A.C., P.H. Wise, and D.V. Kleinman. 1991. “Hispanic Health: Time for Data, Time for Action.” *Journal of the American Medical Association* 265(2):253–55.

Palloni, A. and E. Arias. 2003. “A Re-Examination of the Hispanic Mortality Paradox.” Working Paper No. 2003-01, Center for Demography and Ecology, University of Wisconsin–Madison.

Palloni, A. and J. Morenoff. 2001. “Interpreting the Paradoxical in the ‘Hispanic Paradox’: Demographic and Epidemiological Approaches.” Pp 140–74 in *Population Health and Aging,* edited by M. Weinstein, A. Hermalin, and M. Stoto. New York: New York Academy of Sciences.

Preston, S.H., I.T. Elo, M.E. Hill, and I. Rosenwaike. 2003. *The Demography of African Americans, 1930–1990*. Boston: Kluwer Academic.

Raftery, A. 1996. “Bayesian Model Selection in Social Research.” Pp. 11–195 in *Sociological Methodology,* edited by P.V. Mardsen. Cambridge, MA: Blackwell.

Rogers, R.G., R.A. Hummer, C.B. Nam, and K. Peters. 1996. “Demographic, Socioeconomic, and Behavioral Factors Affecting Ethnic Mortality by Cause.” *Social Forces* 74:1419–38.

Rosenberg, H.M., J.D. Maurer, P.D. Sorlie, N.J. Johnson, M.F. MacDorman, D.L. Hoyert, J.F. Spitler, and C. Scott. 1999. “Quality of Death Rates by Race and Hispanic Origin: A Summary of Current Research, 1999.” *Vital and Health Statistics*, Series 2 (128). Hyattsville, MD: National Center for Health Statistics.

Rosenwaike, I. 1987. "Mortality Differentials Among Persons Born in Cuba, Mexico, and Puerto Rico Residing in the United States, 1979–81." *American Journal of Public Health* 77:603–606.

———. 1991. "Mortality Experience of Hispanic Populations." Pp. 3–14 in *Mortality of Hispanic Populations,* edited by I. Rosenwaike. New York: Greenwood Press.

Rosenwaike, I. and S.H. Preston. 1983. "Age Overstatement and Puerto Rican Longevity." *Human Biology* 56:503–25.

Sandefur, G., M.E. Campbell, and J. Eggerling-Boeck. 2002. "Racial and Ethnic Identity, Official Classifications, and Health Disparities." Discussion paper for the Conference on Racial and Ethnic Disparities in Health, National Academy of Science.

Singh, G.K. and M. Siahpush. 2001. "All-Cause and Cause-Specific Mortality of Immigrants and Native Born in the United States." *American Journal of Public Health* 91:392–99.

———. 2002. "Ethnic-Immigrant Differentials in Health Behaviors, Morbidity, and Cause-Specific Mortality in the United States: An Analysis of Two National Data Bases." *Human Biology* 74:83–109.

Soldo, B., R. Wong, and A. Palloni. 2002. "Migrant Health Selection: Evidence From Mexico and the U.S." Paper presented at the annual meeting of the Population Association of America, Atlanta, GA, May 9–11.

Sorlie, P.D., E. Backlund, N.J. Johnson, and E. Rogot. 1993. "Mortality by Hispanic Status in the United States." *Journal of the American Medical Association* 270(20):2464–68.

Sorlie, P.D., E. Backlund, and J. Keller. 1995. "U.S. Mortality by Economic, Demographic, and Social Characteristics: The National Longitudinal Mortality Study." *American Journal of Public Health* 85:949–56.