FEDERAL RESERVE BANK OF SAN FRANCISCO

WORKING PAPER SERIES

Relative Status and Well-Being: Evidence from U.S. Suicide Deaths

Mary C. Daly
Federal Reserve Bank of San Francisco

Daniel J. Wilson
Federal Reserve Bank of San Francisco

Norman J. Johnson
U.S. Census Bureau

October 2008

Working Paper 2007-12
http://www.frbsf.org/publications/economics/papers/2007/wp07-12bk.pdf

Relative Status and Well-Being:
Evidence from U.S. Suicide Deaths

Mary C. Daly[a], Daniel J. Wilson[a], and Norman J. Johnson[b]

First Draft: May 2006
Current Draft: Oct. 20, 2008

Please cite as:
Daly, Mary C., Daniel J. Wilson, and Norman J. Johnson. "Relative Status and Well-Being: Evidence from U.S. Suicide Deaths." Federal Reserve Bank of San Francisco Working Paper #2007- 12, 2007.

[a] Federal Reserve Bank of San Francisco
[b] U.S. Census Bureau

Corresponding author's email: daniel.wilson@sf.frb.org

We gratefully acknowledge research funding provided by the National Institute of Aging. This paper benefited from helpful comments from Marianne Bitler, Sandy Black, David Card, Raj Chetty, Andrew Clark, Dora Costa, John Ham, Doug Miller, Andrew Oswald, Betsey Stevenson, Justin Wolfers, and seminar participants at SOLE 2006, UC Berkeley, Federal Reserve Bank of New York, and U. of Nevada-Reno. We thank Ann Lucas, Jeremy Gerst, Charles Notzon and Eric Backlund for excellent research assistance. Research results and conclusions expressed are those of the authors and do not necessarily indicate concurrence by the National Institute of Aging, the Bureau of the Census, the Federal Reserve Bank of San Francisco, or the Federal Reserve System.

Relative Status and Well-Being: Evidence from U.S. Suicide Deaths

Abstract:

This paper empirically assesses the importance of interpersonal income comparisons using individual level data on suicide deaths in the United States. We model suicide as a choice variable, conditional on exogenous risk factors, reflecting an individual's assessment of current and expected future utility. Our empirical analysis considers whether suicide risk is systematically related to the income of others, holding own income and other individual factors fixed. We estimate proportional hazards and probit models of the suicide hazard using two separate and independent data sets: (1) the National Longitudinal Mortality Study and (2) the National Center for Health Statistics' Mortality Detail Files combined with the 5 percent Public Use Micro Sample of the 1990 decennial census. Results from both data sources show that, controlling for own income and individual characteristics, individual suicide risk rises with reference group income. This result holds for reference groups defined broadly, such as by county, and more narrowly by county and one demographic marker (e.g., age, sex, race). These findings are robust to alternative specifications and cannot be explained by geographic variation in cost of living, access to emergency medical care, mismeasurement of deaths by suicide, or by bias due to endogeneity of own income. Our results confirm findings using self-reported happiness data and are consistent with models of utility featuring "external habit" or "Keeping Up with the Joneses" preferences.

Keywords: Relative income, interpersonal comparisons, interdependent preferences, suicide, happiness, Keeping Up with the Joneses.

JEL Codes: I31, D6, H0, J0

Relative Status and Well-Being: Evidence from U.S. Suicide Deaths

I. Introduction

Despite a general awareness that we are not myopic in our personal evaluations and a burgeoning literature confirming this in the data, the idea that individual preferences are interdependent has been slow to diffuse into mainstream theory. One reason for the slow adoption of these ideas is that the data used to illustrate the presence and importance of interpersonal comparisons is itself suspect and the subject of considerable concern. Research on relative comparisons has mostly been limited to two types of data: data generated from classroom/laboratory experiments and the results of subjective surveys of happiness or life satisfaction. While innovative and useful, these types of data are subject to two important criticisms: experiments, by their nature, are contrived and typically limited to very small samples; and, self-reported happiness surveys, while capturing much larger samples, elicit responses that are subjective and may be difficult to compare across individuals and over time. These criticisms of the data have limited the widespread acceptance of research findings based on them.

In this paper, we use an alternative measure of utility (disutility), suicide deaths, to examine the importance of relative status. We argue that suicide data have the advantage of directly measuring, through revealed preference, the assessment of utility (happiness) previous research has tried to uncover with self-reported happiness surveys or experiments. Treating suicide as a choice variable regarding current life satisfaction and assessed value of future life, we develop a framework showing how suicide outcomes can be used to explore the importance of relative status in standard models of utility. Based on this framework we write down a reduced form model of suicide risk in terms of own and others' income and estimate it using data from two independent data sources: (1) the National Longitudinal Mortality Study (NLMS) and (2) data from publicly-available death certificates (from the National Center for Health Statistics) combined with the 5 percent Public Use

Micro Sample (PUMS) of the 1990 decennial census. Our results strongly support the idea that individual utility is affected by relative income. In repeated regressions we find that local area (county) median income, holding own income constant, has a significant and positive effect on suicide risk. We show that this result is robust to alternative specifications of the empirical model and that our interpretation that the result supports the importance of relative income is robust to potential alternative explanations. By direct assessment, we conclude that our results cannot be explained by geographic variation in the cost of living, access to emergency medical care, or errors in suicide reporting. In addition, we argue that while there are numerous reasons why county income could *reduce* individual suicide risk (e.g., better mental health care services in higher income counties, endogenous mobility of individuals to counties where their relative income is higher, and county income shocks that are correlated with unobserved non-income shocks (as suggested in Luttmer 2005, for example)), these would result in underestimation rather than overestimation of our main effects. As such, we conclude that our result is consistent with relative income mattering for individual utility.

Having established the robustness of our baseline result and its consistency with previous research from subjective well-being and experimental data, we exploit the richness of our data and expand the empirical literature on interpersonal comparisons along two dimensions. First, we consider whether the relative income effect holds for individuals regardless of their own income level. We find that, controlling for own income, suicide risk rises with median county income both for high-income and low-income individuals, although the effect appears to be somewhat larger for the latter. Second, we consider whether the relevant reference group for interpersonal comparisons is, in fact, narrower or broader than one's local geographic area, defined by county. The results point to age, in addition to local area, being a particularly relevant factor. In contrast, race does not appear to be particularly relevant. We also find that state is too geographically broad to be a relevant reference group: median state income has no discernible effect on individual suicide risk after controlling for one's own income.

The remainder of the paper is organized as follows. In Section 2 we review the empirical work on relative income and utility and discuss how information on suicide fits into and expands the literature. We lay out our theoretical framework and motivate our empirical strategy in Section 3. The data sets we construct and use are described in Section 4. In Section 5, we present our main results and assess their robustness. A summary of our findings and the path for future work are laid out in Section 6.

2. Previous Research

2.1 Existing Literature

Following early recognition of the importance of relative comparisons by Adam Smith, several economists have composed fuller treatments of the issue, including Veblen (1899), Duesenberry (1949), Easterlin (1974), Abel (1990), Galí (1994), Kahneman and Tversky (1996), Frank (2000), Becker and Rayo (2005), and others. These models of interdependent preferences generally posit that individuals care about their own socioeconomic status (generally defined by income, consumption, or wealth), and that of others. A growing empirical literature on the subject has found evidence consistent with this view. Empirical investigations generally can be grouped into two types. The first set consists of controlled experiments contrived to elicit participants' reactions to imposed hierarchies. In these experiments, performed on human and primate subjects, researchers have looked for the subjects' negative reactions to the presence of a hierarchy, i.e., "inequality aversion," and for reactions to subjects' relative placement within a hierarchy, i.e., "interdependent preferences" (Engelmann and Strobel 2004; Brosnan and deWaal 2003; Alpizar, Carlsson, and Johansen-Stenman 2005). Although such experiments consistently find that inequality and relative income matter, the relatively small sample sizes and artificial environments of these experiments make their results difficult to generalize. Moreover, their contrived nature frequently makes it difficult to distinguish inequality aversion from relative income concerns.

A second vein of the literature on interpersonal income comparisons comes from research on responses to questions from subjective well-being (happiness and/or life satisfaction) surveys. A number of researchers have used the responses from these surveys to study the extent to which self-reported happiness or satisfaction is correlated with relative position, holding other factors such as own income constant.[1] For example, Clark and Oswald (1996) use data on 5,000 British workers to investigate whether worker satisfaction rates are inversely related to relative wages. A similar examination is done in Brown, et al. (2005), focusing on relative rankings of workers' wages rather than the relative wage ratio. Both studies find evidence that relative income matters to self-reported satisfaction. Along the same lines, several papers have expanded the potential reference group to which individuals are compared by combining individual data on happiness and income with variables on local, regional, and national income (Helliwell 2002; Luttmer 2005; Tomes 1986; and Ferrer-i-Carbonell 2005). In general, these papers have found empirical support for the hypothesis of interpersonal income comparisons.

Still, serious concerns have been raised about the quality of data on self-reported happiness (see, e.g., Brekke 1997 and Osmani 1993; see Bertrand and Mullainathan 2001 for a broader critique of subjective survey data). These concerns involve language ambiguities (respondents may not all agree on the exact meaning of terms like "happiness" and "life satisfaction"), scale comparability (one person's "very satisfied" may be higher, lower, or equal to another person's

[1]There also is a recent cross-national literature using surveys of happiness. These studies compare average reported happiness to average income across countries. They generally find little correlation (Di Tella, MacCulloch, Oswald 2004; Alesina, Di Tella, and MacCulloch 2001; Easterlin 1973, 1995; Oswald 1997), though an exception is Stevenson and Wolfers (2008) who find strong evidence of a positive correlation.

“satisfied”), ambiguity regarding the time period over which respondents base their answers, respondent candidness, and the difficulty of drawing cardinal inferences from ordinal survey responses. In addition, Diamond (2008) argues that happiness data may be inappropriate for answering the relative income question in particular since the question itself could be a relative one.[2]

Although the results of these two types of studies seem to confirm a role for theories of interdependent preferences, concerns about how representative the underlying data are have hindered broader acceptance of the results. The suggestive findings coupled with concerns about experiments and self-reported measures of happiness suggest that additional methods of addressing the role of relative income are needed.

2.2 Suicide Data as an Alternative

Suicide data provide an alternative measure of happiness (unhappiness) with several advantages over experiments or surveys of happiness.[3] First, suicide can be thought of as a revealed choice made by individuals who have examined the value of continuing to live versus exiting.[4] In studies of consumer choice, using observed choices to infer preferences has long been considered

[2]“How should we interpret answers to the question ‘How happy are you these days?’.... If people answer whether they are satisfied with their lives in terms of their perceived relative position in happiness, that does not necessarily mean that happiness is based on relative position, rather that the question being answered by the respondent is a relative happiness question…. Some exploration has been done of the impact on reported happiness of the...incomes of neighbors. But such studies may not shed light on the question of how much well-being depends on one's relative standing and how much the respondent looks to relative standing in order to answer the survey question.”

[3]As Oswald (1997) puts it, “Suicides represent choices in response to (un)happiness that are intrinsically more compelling than replies made to happiness survey questions, and data that, by their nature, cannot be generated in a laboratory experiment.”

[4]We recognize that the actual choice may be suicide attempt rather than completion. However, data on attempts are quite limited and, moreover, a large share of attempts may reflect “cries for help” rather than true attempts to commit suicide.

preferable to relying on individual self-reports of preferences. Second, suicide data are comparably measured across individuals and regions and over time. Third, in the United States, data on suicides are publicly available and complete, covering the universe of reported suicides by year.[5]

Despite these advantages, suicide data obviously are not perfect. Suicide victims presumably are at the extreme tail of the distribution of life satisfaction over the population, and their preferences may not reflect the preferences of the non-suicide population. We address this concern by comparing our empirical results to those obtained by studies using subjective survey data. Our results are consistent with these previous studies, reinforcing the reliability of inferences based on suicide data. It also is possible that suicide decisions are largely idiosyncratic and unrelated, systematically, to the variables that affect happiness or life satisfaction. While this concern cannot be eliminated *a priori*, if it is binding then we should find no correlation between relative income and suicide risk – a non-finding.

Although frequently regarded as an idiosyncratic act, there is a long history in sociology and economics of relating suicide patterns to more universal social phenomena. The most complete example of such efforts is Durkheim's detailed empirical study of suicide. Durkheim's (1951) Suicide: A Study in Sociology was a careful attempt to analyze the societal influences that effect suicidal behavior and unhappiness more generally. Less comprehensive treatments in economics have nonetheless also treated suicide as a potential social phenomenon, affected by both societal and individual factors. Examples of this work include Hamermesh and Soss (1974), who develop an economic theory of suicide and, using cross-country and cross-state data, find that suicide risk is significantly related to unemployment and decreases in permanent income. More recently, Ruhm

---

[5]Reported suicides may undercount all true suicides; many experts believe that a significant share of true suicides are misclassified as accidents or "undetermined injuries" (see Moyer, Boyle, and Pollock 1989; Rockett and Smith 1999; and Mohler and Earls 2001). We address this possibility in our empirical analysis.

(2000) considers suicide as one of several causes of death and finds that, unlike other negative health outcomes that decline during times of recession, suicide risk is either increased or unaffected. In other work, Helliwell (2004) investigates the empirical association between subjective well-being and suicide rates using cross-country panel data and finds a strong negative relationship.[6] In a related survey article on happiness and economic factors, Oswald (1997) notes that many variables positively (negatively) associated with reported happiness are negatively (positively) associated with suicide risk.[7] To our knowledge, though, we are the first to use information on suicide risk to study the existence and nature of interpersonal comparison.

3. Conceptual Framework and Empirical Strategy

A natural question to ask is whether the concept of suicide and its measured occurrence can be used to understand the parameters of the utility function for the general population. In this section we argue that it can and provide a framework for understanding the relationship between suicide and population happiness. We then review our empirical strategy for gleaning information about preferences from analysis of suicide data.

3.1 Suicide and Happiness

We begin with the idea that individuals assess their current and expected future utility, and in so doing compare the value of future life relative to the current value of exit (suicide). For the

---

[6]Similarly, Koivumaa-Honkanen, et al. (2001) find that individual self-reports of life satisfaction have significant predictive power for suicide over the subsequent 20 years.

[7]Other recent examples of economists trying to explain suicide behavior include Cutler, Glaeser, and Norberg (2000), Brainerd (2001), Marcotte (2003), Stevenson and Wolfers (2000), Chuang and Huang (1997), Huang (1996), Kimenyi and Shughart (1986), Hamermesh (1974), and Schapiro and Ahlburg (1982-83). There have also been a number of recent studies in the psychiatry and public health literatures exploring the empirical links between suicide and socioeconomic factors (see, e.g., Blakely et al. 2003, Lewis and Sloggett 1998, and Kposawa 2001).

vast majority of the population, the future value never nears the low threshold defining exit but rather determines the extent to which someone experiences being happy or not so happy. For the marginal person, however, the threshold is binding and the evaluation ends in suicide. Under this reasoning, we can use data on suicide deaths to directly test hypotheses related to the utility function, such as the extent to which neighbors' income or socioeconomic status affects utility.

To get a clearer understanding of how we relate suicide to population happiness it is useful to consider the schematic in Figure 1. The figure shows the happiness continuums for two individuals. A and B, as well as their suicide thresholds θ_i. As the schematic illustrates, our maintained hypothesis is that factors affecting utility (X_i) have the same marginal effect on suicide risk as they do on happiness, but that thresholds for suicide differ across individuals.[8] That is, suicide victims and the general population have the same β's but different θ's. This means that uncovering the marginal effects of variables on suicide informs us about how these variables affect happiness for the rest of the population.

Absent empirical work, we cannot know whether individuals who commit suicide differ only in θ or in both θ and β. In a cross-validation exercise using subjective well-being data and suicide, Daly and Wilson (2008) find evidence supporting the idea that β's are the same between those who commit suicide and those who do not. Specifically, they find that the relative risks of suicide along a host of dimensions (such education, income, age, marital status, and employment status) closely match the relative risks of reported unhappiness. Based on these results we proceed as if the model in Figure 1 is reasonable. However, following our main results, we conduct several checks to insure that this hypothesis is consistent with the evidence.

Based on this intuition, we turn to modeling suicide. Our formal model is described in Appendix A. The basic idea is that suicide is an outcome of an individual's dynamic optimization

---

[8]An alternative schematic, not shown, would illustrate the case where the θ_{is} —are the same

problem weighing the present discounted value of current and expected future lifetime utility against the value of exiting life right now. Individuals differ in their inherent levels of happiness (set points) or suicide thresholds—θ_i—and in their observable circumstances, $\mathbf{X}_i$. The linear preferences of the general population—the β's—can be estimated by observing how changes in **X** affect the propensity to commit suicide, under the identifying assumption that the β's are the same for suicide victims as for the rest of the population. While the vast majority of the population never commits suicide, since the future value never nears the low threshold defining exit, the determinants for suicide and being happy or not so happy are the same. For the marginal person, however, the threshold is binding and the evaluation ends in suicide. Under this reasoning, we can use data on suicide deaths to directly test hypotheses related to the utility function, such as the extent to which neighbors' income or socioeconomic status affects utility.

3.2 Empirical Strategy

As illustrated with the schematic discussed above and the model in Appendix A, the probability that an individual commits suicide is simply the probability that his/her θ_i is below the threshold level $\overline{\theta}(\mathbf{x}_{it})$. Letting G denote the CDF of θ_i, then

$\Pr\left[\theta_i < \overline{\theta}(\mathbf{x}_{it})\right] = G\left[\overline{\theta}(\mathbf{x}_{it})\right] = F(\mathbf{x}_{it})$. Parameterizing $F(\mathbf{x}_{it})$, a researcher equipped with individual level data on suicide decisions and the variables in $\mathbf{x}_{it}$ can estimate the average effect of any variable in $\mathbf{x}_{it}$ on the likelihood of suicide, and thus infer the effect of the variable on utility.[9] We estimate models based on two alternative parameterizations. The first is the proportional hazards model,

---

and the the β's are different.

[9]Note the disturbance term in the estimation can be interpreted as an estimate of θ_i.

$$h(\tau \mid \mathbf{x}_{i,0}) \equiv \frac{f(\tau \mid \mathbf{x}_{i,0})}{1 - F(\tau \mid \mathbf{x}_{i,0})} = \psi(\tau)\exp\left(\mathbf{x}_{i,0}\beta\right), \tag{1}$$

where $h()$ is the suicide hazard at duration τ from when $\mathbf{x}_{i,0}$ is observed and $f()$ is the PDF corresponding to the CDF, $F()$. The second parameterization is the probit model,

$$F(\mathbf{x}_i) = \Phi(\mathbf{x}_i\beta), \tag{2}$$

where $\Phi()$ is the standard normal CDF.

Our objective is to estimate the effects of own income and reference-group income on the probability of suicide. Given the longitudinal nature of the NLMS data and the cross-sectional nature of the MDF-PUMS data, we estimate a Cox proportional hazards model with the former and a probit model with the latter. The probit model essentially regresses individual suicide probability on a function of own income, reference-group income, and a set of control variables. Similarly, using the NLMS data, we estimate Cox proportional hazards models of suicide risk–i.e., the hazard rate of suicide in a given period–as an exponential function of own income, reference-group income, and a set of controls. The Cox proportional hazards models allows us to easily control for the duration that sample members are observed (exposed) and for right-censoring, i.e., the fact that suicide is not observed for most individuals due to either nonsuicide death or the end of the follow-up window.

4. Data and Variable Specification

This paper uses two alternative individual-level data sets, each with their own relative advantages, to analyze the relationship between relative income and suicide. The first data set is the National Longitudinal Mortality Study (NLMS) augmented with data on county and state income from the U.S. Census Bureau. The second data set we constructed by combining the Mortality Detail Files (MDF) for years 1989-1992, with data from the 1990 5% Public Use Micro Sample

(PUMS). We will refer to this data set as the MDF-PUMS data. The MDF-PUMS data have the advantage of containing a very large number of observations on suicide victims (as well as on the general population), whereas the NLMS data have a much smaller sample of suicide and nonsuicide records from which to draw inferences. On the other hand, the mortality records in the MDF data do not include income and do not identify county of residence for sparsely populated counties, whereas the NLMS contains actual, reported income and has no limitations on geographic coverage. We choose to restrict both data sets to working-age adults (20-64), for whom relative income concerns are likely to be most relevant. Each of these two data sets is described in detail below.

4.1 NLMS

The NLMS is a confidential, restricted-use database developed and maintained by the U.S. Census Bureau to facilitate research on the effects of demographic and socioeconomic factors on mortality (see U.S. Bureau of the Census 2005). It has been used extensively by epidemiologists and public health experts in recent years, for example to study cancer and heart disease, but it has not previously been used by economists (to our knowledge). The NLMS consists of a set of cohort files, primarily from Current Population Surveys (CPS), matched to the National Death Index (NDI), a national database containing the universe of U.S. death certificates since 1979. The cohort files included in our analysis–those with sufficient information on income–are all March CPS files from 1979 to 1998, plus CPS files for February 1978, April 1980, August 1980, and December 1980. At the time of this writing, the mortality follow-up (i.e., the matching to the NDI) from the cohort files covered deaths occurring from January 1, 1979, through December 31, 1998. The matching process appends to individual CPS records (1) whether the person has died within the follow-up period, (2) date of death (if deceased), and (3) cause of death (if deceased).

For our analysis, we restrict our sample to non-Hispanic working-age adults. We exclude

Hispanics because of definitional changes in the Hispanic status variable over time and because of concern that a nontrivial share of Hispanic CPS respondents may have moved out of the United States prior to the end of the follow-up period, in which case their deaths would not be observed. Given the richness of information available from death certificates, we are able to identify whether a suicide victim had a diagnosed severe mental illness.[10] As our goal is to infer preferences of the general population, we exclude these (160) suicide cases from our primary sample.[11] The final data set, after excluding a relatively small number of records with missing values for key variables, contains 957,934 individual records, including 74,786 nonsuicide deaths and 1,401 suicide deaths within the follow-up period (the remainder were still alive as of December 31, 1998).

We merge onto the NLMS data on a number of geographic aggregates, most notably mean family income by county-year. The construction of these geographic aggregate variables is described in Appendix B.

4.2 MDF-PUMS

The public use MDF, compiled by the National Center for Health Statistics and available from the Inter-university Consortium for Political and Social Research (ICPSR), for a given year are essentially the data from all death certificates recorded in the United States in that year (see U.S. Department of Health and Human Services 1992).[12] For the years 1989-1992, we extract the records where suicide is the cause of death (i.e., International Classification of Death, Rev. 9 (ICD9)

[10]Known mental illnesses are listed as "contributing factors" on the death certificate.

[11]We note, however, that results based on the full sample are quite similar to those for this restricted sample though the estimated coefficients generally are closer to zero (as one would expect if severely mentally ill individuals place less weight on socioeconomic factors in the utility function and/or simply behave less "rationally").

[12]For later years, these data are called Multiple Cause of Death files. These data are also sometimes referred to as the Mortality Detail Files.

codes E950-E959) and combine them with the individual records from the PUMS 5 percent sample of the 1990 decennial census (Ruggles et al. 2004), which we treat as nonsuicide observations. We extract suicides for years other than 1990 to maximize the number of suicide observations, given that suicide is a relatively infrequent event. In our empirical analysis, we adjust for the fact that nonsuicides are undersampled relative to suicides in our data, both because we include four years of suicide records (versus one year from PUMS) and because the nonsuicides are a 5 percent sample of the overall population.

The variables jointly available in the MDF and the PUMS are age, race, sex, county of residence, marital status, education, and Hispanic status. Income, on the other hand, is not recorded on death certificates. We therefore estimate income by matching suicide records in the MDF to individuals or groups of individuals in the PUMS data, where income is available. The matching procedure works as follows: (1) for each suicide record, find all matching observations in the PUMS, with its roughly 7 million records, matching on county, age, race, sex, Hispanic status, education, and marital status; (2) calculate average family income for this matching cell; and (3) assign this average income to the suicide observation. This procedure provides a reasonably accurate estimate of income: over the 7,202,093 working-age observations in PUMS, county, age, race, sex, Hispanic status, education, and marital status jointly explain 24 percent of the individual level variation in family income.[13] A variance decomposition (not shown) reveals that county, education, and marital status (in decreasing importance) have the greatest explanatory power, together accounting for 16 percent of variation.

With this matching procedure, we are able to estimate family income for 57 percent of U.S.

[13]Including occupation and industry in the income estimation would modestly improve the model fit to 28 percent. However, less than half of the suicide records report occupation and industry (as many states do not include them on death certificates). Therefore, we omit these variables from the matching procedure.

working-age suicide records from 1989-1992, totaling 50,328 suicides.[14] We use the same matching procedure to generate an analogous predicted income variable for the nonsuicide records; this is the "own" income variable used in our regression analyses. The final data set has 4,360,747 observations.

We additionally use the PUMS data to construct several control variables at the county-age-race-sex-education-Hispanic-marital status cell level: share of cell that owns a home, share of cell that are Vietnam veterans (found in previous research to be a correlate with suicide), and the unemployment rate within the cell. Another control we include in our analysis is a state-level measure of firearm availability from Miller et al. (2002); specifically, it is the share of suicides in the state that are committed via firearm. (Using the share of homicides committed via firearm yields similar results.)

4.3 Descriptive Statistics: Suicide Risk and Model Variables

National statistics show that the U.S. suicide rate has been relatively constant since 1950, averaging about 12 per 100,000 persons (see WHO 2005).[15] Table 1 reports suicide risk overall and by our model variables for the NLMS and MDF-PUMS samples. Recall that both samples exclude Hispanics and cover only working-age adults. The overall suicide rates in the NLMS and MDF-PUMS are quite similar to each other, at approximately 13 per 100,000, and are comparable to the

[14]The main constraining factors here in terms of coverage are county of residence and education. Education is simply unknown or unreported on many death certificates. For confidentiality reasons, county of residence (or occurrence) is not identified on the public-use MDF data if the county has a population below 100,000. This occurs for roughly a quarter of U.S. counties in 1990, covering slightly more than a quarter of all suicides. It should also be noted that some death records include occupation and industry of the deceased, but not enough records contain this information for us to include these variables usefully in our matching procedure.

[15]From 1950 to 2000, the overall U.S. suicide rate has fluctuated within the narrow range of 10.4 to 13.5 per 100,000. The typical rate for the working-age adult population is somewhat higher, around 12 to 15 per 100,000.

national statistics. Furthermore, national data indicate considerable variation in suicide risk by gender, age, and race. These patterns are mirrored in the NLMS and MDF-PUMS samples. For example, suicide rates are far higher for males than for females and higher for whites than for other races. Suicide rates decline slightly with age according to the MDF-PUMS while having no clear age trend in the NLMS sample, which may simply be due to the relatively small sample size of the NLMS. In both samples, married individuals have a lower suicide rate on average relative to those who are single/never married or divorced/separated. Suicide rates generally fall, though not monotonically, with educational attainment. Although rudimentary, these categorical suicide rates suggest that the two data sources used in our analysis produce patterns consistent with the stylized facts regarding suicide reported in the epidemiology/public health, psychology, and sociology literatures.

The key variables in our analysis are own and reference group income. To assess the extent to which preferences of the general population can be inferred from the revealed preferences of suicide victims, it is helpful to first compare these two populations along the key dimension of income. Figures 2 and 3 plot the distribution of predicted family income for working-age suicide victims in our two samples against the income distribution for the general U.S. working-age population.[16] Figure 2 shows the distributions of reported family income (adjusted to 1990 dollars) for the total sample and for the subset of those who eventually commit suicide, according to the NLMS data. Note that the NLMS data are survey reports reflecting income at the time the individual was surveyed rather than income at the time the suicide was committed. The income distribution of suicide victims is slightly left of that for the general population. That said, the bulk of the suicide population has income in the middle range of the distribution. We take this as

[16]Recall that both the suicide and general populations in the MDF-PUMS sample exclude individuals from counties with population under 100,000, since such counties are not identified in the data for confidentiality reasons.

supporting evidence for the notion that suicide victims are broadly representative of the general population, at least in terms of income. This will aid us when we offer an interpretation for our later findings.

Figure 3 reports income figures for the MDF-PUMS sample; the figure shows the distribution of estimated family income (estimated as described in Section 4 above) of suicide victims compared to estimated family income of the general population. The distributions suggest that the modal suicide victim sits slightly to the left of the modal member of the general population, but overall the two distributions are quite similar. Importantly, there is little difference in the lower tail of the income distribution and overall the shapes for the two populations are roughly similar.[17] The fact that the MDF-PUMS data show a pattern similar to the NLMS data suggests that our estimated income data in the MDF-PUMS data set are reasonably accurate.

Turning to county income, suicide risk has a strong negative correlation with county income. One can see this in Figure 4 which shows a scatterplot of county suicide rates (from the MDF) and county income per family in 1990. Each circle in the plot represents a single county and the size of the circle is proportional to the county's population. The unweighted correlation is −0.07 and the population-weighted correlation is −0.29; both are significant at well below the 1% level. Note we also have confirmed that this negative (unconditional) correlation between suicide risk and county income is present in the NLMS sample with a simple proportional hazards model of suicide risk regressed on county income alone (results available upon request). Thus, it is clear that the positive effect of county income on suicide risk that we find later in our multivariate results is not what one would expect *a priori*.

Descriptive statistics for other model variables are reported in Tables C1 (NLMS) and C2 (MDF-PUMS) of Appendix C. Again, the key variables in our analysis are of similar magnitudes

[17] We also did this matching using education alone and obtained similar results. Full details of both estimation strategies are available from the authors upon request.

and have similar patterns in both data sets.

5. Results

5.1 NLMS Baseline Regression Results

In this subsection we describe the results from estimating the Cox proportional hazards model, equation (1) above, using the NLMS sample. We use Cox proportional hazards models to accommodate the unbalanced nature of the NLMS data. The NLMS data provides socioeconomic and demographic information on individuals at the date of their initial CPS response (period 0) as well as time elapsed until death. Thus records are identified in terms of duration from the original interview rather than in chronological time. The proportional hazards model allows us to characterize the suicide hazard (probability of suicide after τ periods given it has not already occurred) over the interval from 0 to T, where T is the maximum duration in the sample, conditional on individual covariates recorded at period 0. In the NLMS, T is 7,633 days, which is the difference between Dec. 31, 1998, the end of the NLMS follow-up window, and Feb. 1, 1978, the date of the earliest CPS response in the sample. Note that the vast majority of observations (individuals) are censored. Observations can be left-censored due either to non-suicide death prior to the end of the follow-up period or to participating in a later CPS survey than February 1978. Observations can be right-censored due to the individual still being alive at the end of the follow-up period. The estimation procedure accounts for both left- and right-censoring.

The estimated coefficients and associated p-values for our baseline models are reported in Table 2. The p-values for these and all other regressions reported in the paper are based on standard errors that are robust to heteroskedasticity and clustering within county. Before turning our attention to the estimated effects of income variables, let us first briefly discuss the effects of the various control variables included in the model. The bottom portion (panel C) of Table 2 confirms that sociodemographic factors are important determinants of suicide risk.[18] Consistent with the raw categorical suicide rates in Table 1, being female or nonwhite lowers suicide risk, while being divorced or widowed, separated, or never married raises suicide risk (relative to being married). Veterans are found to be more likely to commit suicide than nonveterans.

There is little evidence of a conditional age profile to suicide risk, though the point estimates suggest perhaps a weak inverted-U age profile. Controlling for these other factors as well as income, educational attainment does not have a separate statistically significant influence on suicide risk, though the point estimates suggest a mildly negative relationship between education and suicide risk. Labor market status, however, does have an important additional influence on suicide risk: being unemployed or out of the labor force, for any reason, raises suicide risk relative to being employed.[19] Specifically, those that are out of the labor force because they are unable to work have the highest suicide risk, followed by unemployed members of the labor force, retired persons, and those that are employed but not currently working (e.g., persons on furlough). In terms of magnitude, the estimated coefficient on unemployment of 0.577 implies a hazard ratio of 1.78 ($e^{0.577}$), meaning that suicide risk (within the subsequent 12.7 years, the average exposure time) for unemployed persons is 78% higher than it is for those who are employed and working.

The magnitude and statistical significance of the coefficients on the control variables is little affected by which income variables are included in the model. Thus, for the remainder of the discussion, we will limit ourselves to discussing just the coefficients on the income variables.

The key variables in our analysis–own income and relative income–are shown in the upper

[18]In unreported regressions, we reestimated the models that follow including year dummies indicating the year of the CPS. The estimated coefficients on the model variables are quite similar to those reported below, though the standard errors are moderately higher. Results are available upon request.

[19]The high relative risk of suicide for unemployed individuals has been found previously using similar data (Kposawa 2001, Blakely, et al. 2003).

portion of the table. Column 1 focuses on the importance of own family income, measured linearly in logs. Own income is statistically significant and negative, implying that higher own income lowers suicide risk. The coefficient on log own income of -0.080 suggests that a 10% higher income is associated with 0.8% lower suicide risk. Column 2, however, shows that this average effect masks the fact that the effect of income on suicide varies meaningfully and nonlinearly across the income distribution. The results indicate that individuals with family incomes below $20,000 in 1990 dollars (which, by way of reference, is equivalent to about $31,000 in 2006 dollars, according to the CPI-U) are significantly more likely to commit suicide than those with incomes above $60,000 (≈$92,500 in 2006 dollars). In contrast, for those with incomes over $20,000, own income has no significant effect on suicide risk.[20] The point estimates of the coefficients on the categorical income variables imply hazards ratios of 1.52, 1.47, 1.12, and 0.99, respectively, for income categories [0,10K], [10K, 20K], [20K, 40K], and [40K, 60K]. A hazard ratio of 1.52, for instance, means that an individual with family income less than $10,000 (in 1990 dollars) is 52 percent more likely to commit suicide than an individual with income above $60,000 (the omitted income category). The hazard ratios decline more or less monotonically toward 1.0 as income approaches the omitted top category (for which the hazard ratio is implicitly 1.0). This pattern is consistent with the standard assumption of diminishing marginal utility of income/consumption. Given the significance of the income gradient in determining suicide risk, we use this model going forward.

The key result for our analysis is shown in Column 3, which displays results of adding reference group income. Following previous work on interpersonal income comparisons, our baseline specification considers county of residence to be one's reference group. The results show

[20]Previous research on the individual effects of own income on suicide is inconclusive. Similar to our finding, Kposawa (2001), using an earlier version of the NLMS, found that in a multivariate regression, suicide risk decreases with income. Lewis and Sloggett (1998) and Blakely et al. (2003), however, using British and New Zealand data, respectively, found no significant effect of income after other determinants of socioeconomic status had been controlled for.

that county income has a positive effect on suicide risk controlling for own income, implying that a loss of relative position leads to a reduction in individual happiness. This finding is consistent with the results of studies using happiness survey data. Our estimated coefficient of 0.306 on log county income implies that, holding own income constant, a 10% higher county income is associated with about a 3% higher suicide hazard relative to the baseline hazard (conditional mean hazard).[21] The final column of Table 2 checks the robustness of our baseline specification against the possibility that the results on county income rather reflect differences in county demographic composition that are correlated with both increased suicide risk and higher county income. Including variables for county population shares by age and race does not alter the qualitative result; in fact, including these shares increases the magnitude of the coefficient on county income to 0.392.

5.2 NLMS Robustness Checks

The results in Table 2 suggest that county income has a positive effect on suicide risk controlling for own income. Before assigning a behavioral interpretation to this result however, there are a number of alternative explanations that must be considered.

Endogenous Income. One potential concern is that own income is endogenous, reflecting unobserved factors, particularly mental health status, that affect both suicide risk and own income.[22] In this case, our baseline results might be biased both for the estimated effect of own income (likely

[21]The proportional hazards function is $h(t) = h(0)e^{\alpha \ln(\bar{y})}e^{X\beta}$, where y is county income and X is a vector of all other model variables. The elasticity of the hazard with respect to county income is then: $\mathrm{dlog}(h(t))/\mathrm{dlog}(\bar{y}) = \alpha \mathrm{dlog}(\bar{y})$. We estimate $\hat{\alpha} = 0.306$.

[22]Tekin and Markowitz (2005) found that suicide ideation has a negative effect on one's wages, suggesting that mental health (illness) may have a positive (negative) effect on income. We also note that when we add the 160 observations from suicide victims for whom a mental illness was identified on their death certificate to our sample (the exclusion of these cases in our main sample is discussed in Section 4.1), the estimated own income gradient becomes steeper, suggesting that severe mental illness is negatively associated with income.

negatively) as well as the estimated effect of county income (and other variables).[23] To account for this possibility, we use a two-stage estimator in which own income is treated as endogenous and instrumented using state of residence dummies (along with all other independent variables included in the model). For simplicity, we depart from our preferred model, which measures income categorically, and estimate a model containing a single continuous own income variable. The results are reported in Table 3. For purposes of comparison, column 1 of the table reports our baseline model, replacing categorical income with a continuous measure. Column 2 shows the results using our two-stage estimator. The two-stage results do suggest the baseline effect of own income was biased, but somewhat surprisingly the direction of the bias is upward (in absolute value) rather than downward as expected; in other words, in the two stage model, own income has a larger (more negative) effect on suicide risk than in the baseline model.[24] Most importantly, the results do not alter the qualitative result on county income: controlling for the endogeneity of own income, county income remains positive and significant. In fact, the coefficient on county income from the two-stage estimation is somewhat larger than in the baseline model.

Cost of Living Differences. Another potential explanation for the positive effect of county income on suicide probability relates to the idea that county income is really a proxy for cost of living, implying that our results reflect that, conditional on nominal own income, individuals are made worse-off by living in areas with a higher costs, especially costs on nontradables such as housing. Given the feasibility of this alternative, we construct several tests to checks its explanatory power. The results of these tests are reported in Table 4; for convenience of comparison, our preferred model from Table 2 is repeated in column one. The first test we employ is to add state

[23]Note that our probit regressions using the MDF-PUMS sample are immune from this potential bias since our own income measure is already a fitted value from a first-stage estimation using PUMS data.

[24]In addition, a Durbin-Wu-Hausman test rejects the exogeneity of own income.

fixed effects to our baseline regression. The logic is that regional differences in cost of living, associated with location, tax structures, etc., will be captured at the state level and pulled out in the state fixed effect. To the extent that these cost of living differences are driving our results, the coefficient on county income should fall or become insignificant. The results in column two shows the opposite occurs: the coefficient on county income rises slightly, from 0.392 in the baseline, to 0.437 when state fixed effects are included. The coefficients on own income are relatively unchanged.

Our next two tests exploit the fact that the cost of housing is likely the most important component of cost of living differences across areas. Column three reports results from adding a county quality-adjusted house price index (described in Appendix B) to our baseline regression. Given the quality adjustment, this index reflects the average cost of land in a county (in a given year) as well as any differences across counties in construction costs. The inclusion of the index drives up the coefficient on county income, while the coefficients on own income remain qualitatively unchanged. The coefficient on the index of housing costs is negative, suggesting that suicide risk is lower in counties with higher housing costs, perhaps because these costs reflect positive area amenities capitalized in local land values.[25] Finally, we use the fact that high house prices have different impacts on owners and renters. For owners, higher house prices increase their cost of living but also contribute to their wealth, which should help lower suicide risk. For renters, high house prices translate solely into higher cost of living, with no offsetting benefits. Thus, if county income simply proxies for cost of living, then the effect of county income on the suicide hazard of renters should be higher than the effect on owners. Column 4 of Table 4 shows the results of interacting county income with a renter dummy variable. We find no significant difference in the

[25]Interestingly, in a regression with both state fixed effects and the county house price index included, the coefficient on the house price index is close to zero and statistically insignificant, suggesting that the variation in cost of living is primarily state level.

effect of county income on renters' suicide hazard relative to owners, refuting the notion that county income is just a proxy for cost of living. Based on these checks we conclude that our results are not picking up the negatives of higher costs of living, but rather are capturing the effects of relative income on suicide risk.

County Income and Mortality. A third potential concern is that our estimated effect of county income on suicide risk reflects spurious correlation associated with some unknown and/or unmeasured relationship between county income and mortality in general or quality of emergency medical care. Moreover, one might be concerned that the universe of reported suicides represents a selected sample of all suicides and one that is not reflective of a more comprehensive measure. The results in Table 5 are designed to address these concerns. The first column simply redisplays our baseline results from Table 2, column 3. The second column repeats the analysis adding deaths from "injuries of undetermined cause" (ICD9 codes E980-E989), which some have argued primarily capture unreported suicides, to the reported suicide records. The results are quite similar to those for reported suicides alone.

The third column of Table 5 reports results using death rates from heart attacks (acute myocardial infarction, ICD9 code 410) as the dependent variable. Our use of heart attack deaths is meant to test whether our results on suicide risk owe to differential quality or access to emergency room care or paramedical care, rather than to reactions to relative income. Research has shown that heart attack deaths are correlated with time to treatment (e.g., proximity to emergency rooms). If our results on suicide are due to unequal access to emergency rooms such that attempted suicides more frequently end in death, then we should see the same pattern for heart attack deaths. This is not the case. Indeed, while the mortality hazard from heart attacks falls monotonically with own income, as with suicide, it also falls with county income, contrary to suicide. Finally, looking at all causes of mortality, our findings concur with the standard result in the literature (see, e.g., Miller and Paxson 2006 and Gerdtham and Johannesson 2004): mortality falls monotonically with own income and is unaffected by relative income.

Based on these results, we conclude that our finding of a positive effect of local area income on suicide, after controlling for own income, reflects a behavioral response to unfavorable interpersonal income comparisons.[26] These individual-level results are consistent with earlier, semi-aggregate results for suicide risk (Daly and Wilson 2006) and with recent empirical analyses using self-reported, subjective well-being survey data (Luttmer 2005).

5.3 NLMS Extensions

We now turn to extensions on our basic specification. We first consider expanding the reference group to the state level.[27] Column 1 of Table 6 again redisplays the baseline results from Table 2, column 3, while column 2 shows the results from replacing county income per family with state income per family. The results show that state income has no significant effect on suicide risk, suggesting that the state is too large to be considered as a relevant reference group. The third column considers whether or not the relative income effect varies over the income distribution. To do so, we interact the categorical income variables with county income. The results are shown in Panel B. While the small sample size limits the statistical power in this regression, the higher point estimates of the interactions involving the lower income categories are suggestive of a stronger

[26]One might worry that our results are reflecting unmeasured correlation between county income and unobserved county characteristics, such as mental health services, that also are correlated with county income. While we cannot rule this out completely, we note that concerns along these lines likely would produce a downward bias on the county income effect. For instance, previous research has shown that psychiatric services are positively correlated with county income (see, e.g., Zimmerman and Bell 2006).

[27]Consideration of reference groups at a finer disaggregation than county is not possible with our NLMS sample due to lack of income data availability over time. We do, however, investigate narrower reference groups below with our MDF-PUMS sample, which requires reference group income data only for 1990, a decennial census year.

effect for those at the bottom of the income distribution than for those at the top.

5.4 MDF-PUMS Regression Results

Despite the strong evidence of interpersonal income comparisons coming out of the NLMS data, one might still be skeptical of this result given the relatively small number of suicides in the NLMS data. To assess whether this result is unique to the NLMS sample, and hence in some way spurious, and to further explore the relevance of alternative reference group definitions, we now consider a second, separate data set we call MDF-PUMS. As described in Section 4.2, the MDF-PUMS data are produced by combining suicide records from death certificate data with individual records from the PUMS 5% sample of the 1990 decennial census. We estimate a probit model (equation (2) above) of the probability of committing suicide as a function of (log) estimated own income, (log) average county income, and various controls. Note that in these models we use a pared down set of sociodemographic controls that does not include the three most important income predictors used in our estimation of family income: county, education, and marital status. Otherwise, due to multicollinearity between these variables and predicted income, there would be little independent variation with which to identify the coefficient of own income.[28] Our strategy in these regressions thus amounts to treating estimated income as a summary statistic for socioeconomic status.

Table 7 shows results for our full non-Hispanic sample. The table reports the estimated marginal effects evaluated at the mean (adjusted for the oversampling of suicides (MDF) relative to nonsuicides (PUMS) in our sample), as well as their corresponding p-values. The p-values are based on standard errors that are robust to heteroskedasticity and clustering within county. To the

[28]We confirmed this point by running a regression equivalent to that in Column 2 of Table 7 but that additionally included education and marital status dummies. As expected, the effects of own income were essentially unidentified (i.e., the standard errors were extremely large).

extent possible we have matched the analysis reported for the NLMS data in the MDF-PUMS sample. The first column shows a baseline regression with our control variables and estimated own income.[29] Note the conditional mean probability of suicide is estimated to be 0.000103, or 10.3 per 100,000, which is similar to the sample suicide rate (unconditional mean probability) of 13.4 per 100,000 reported in Table 1. The coefficients on the control variables are broadly consistent with our NLMS results and with previous research: suicide risk is considerably higher for males and for whites and exhibits an inverted-U age profile.[30] The Vietnam veteran share and the unemployment rate, measured for an individual's county, age, race, sex, Hispanic status, education, and marital status cell group, both have an upward effect on suicide risk, while cell-level homeownership has a downward effect – all consistent with the NLMS results. The firearm availability in one's state (proxied by the share of suicides committed by firearm) is positively associated with suicide risk.

As in the NLMS/proportional hazards estimation, we find that suicide risk is decreasing in own income. Column two reports results allowing own income effects to vary across the income distribution. As in the NLMS data, we find a steep income gradient for own income and suicide risk. Own income changes have a larger effect on suicide risk at the low end than at the high end of the income distribution.

The remainder of Table 7 focuses on the effects of county income. Column three reports our baseline regression of including county income. As in the NLMS, the results for the MDF-PUMS point to a significant effect of county income on suicide risk, holding own income constant. The estimated effect of county income implies that, holding an individual's own income constant, a

[29]In robustness checks not shown here we adjusted a subset of the models for the fact that income is an estimated variable using the technique developed by Murphy and Topel (1985). In each of the cases we tried, the adjustment had a negligible effect and made no material difference in our findings.

[30]It is worth noting that studies using subjective survey data have tended to find that subjective well-being is U-shaped in age (e.g., Blanchflower and Oswald 2004), consistent with the inverted-U age profile for suicide that we find.

10% higher county income is associated with an increase in suicide probability of 7.5%.[31] Recall that the corresponding number from our analysis of the NLMS sample was about 4%. Column four adds state fixed effects to account for potential differences in cost of living across regions. Including state fixed effects does not change the significance or the size of the county income variable. Finally, the last column of Table 7 reports results including the county house price index (discussed above and in Appendix B). Again, and as was the case in our NLMS analysis, county income remains positive and significant, although with some decline in the size of the effect. Overall, the results from the MDF-PUMS to be strikingly similar to those from the NLMS and consistent with the idea that relative income matters for suicide risk (unhappiness).

The final component of our analysis exploits the greater detail and sample size in the MDF-PUMS to consider the importance of different reference groups for relative income comparisons. Table 8 reports results from introducing different reference income values computed over various sub-reference groups. Column one repeats the baseline results from Table 7. The next two columns examine the effect of changing the reference group definition–in this case narrowing it. The results suggest that, while others in one's county or others of the same race in one's county are relevant reference groups, others in the same age range in one's county may be the most relevant reference group. The final column shows that the average state income has a positive but statistically insignificant effect on an individual's suicide risk–again, consistent with the NLMS results–implying that the relevance of others in one's reference group declines with distance and one's effective reference group may be geographically narrower than the state.

[31] This number is based on the following calculation: The coefficient on log county income in Table 7 equals 0.000245, which implies that an increase in log county income of 0.10 is associated with a decline in the suicide probability of 0.10*0.000245 = 0.0000245. The estimated probability of suicide when all regressors equal zero is 0.000328 (see penultimate row of Table 7); 0.0000245/0.000328 = 0.075.

6. Additional Concerns and Future Research

Using individual level data on suicide risk, we find compelling evidence in support of the idea that individuals care not only about their own income but also about the income of others in their local area. This finding is obtained using two separate and independent data sets, suggesting that it is not an artifact of the particular sample design of either data set. Importantly, the finding is robust to alternative specifications and cannot be explained by geographic variation in cost of living, access to emergency medical care, suicide reporting behavior, or by bias due to the endogeneity of own income.

It is also worth noting that other plausible stories of potential bias that we cannot test or rule out with our data, generally imply a *downward* bias on our key county income variable. For instance, previous research has shown that psychiatric services are positively correlated with county income (Zimmerman and Bell 2006). This positive correlation combined with the possibility that the quality of local mental health care negatively affects suicide hazard implies a possible downward bias on county income's effect on suicide. Another possibility is that individuals are mobile and endogenously select their county of residence in response to their income relative to the county's average (assuming, perhaps unrealistically, that individuals can obtain the same income when they move). This would suggest that suicide outcomes underestimate the true relevance of interpersonal income comparisons because individuals are able avoid the negative utility impact of low relative income by simply moving to a location where they have higher relative income. Another possible story is that county income shocks may be correlated with unobserved non-income county shocks that reduce the general well-being of county residents and hence increase suicide risk. For instance, a local plant closing might both reduce average household income in the county and lead to other negative county-wide outcomes (reduced local tax revenues and public services, reduced social capital, etc.) that are unobserved and reduce utility of individuals in the county, hence increasing suicide risk. Luttmer (2005) investigates this possibility in the context of reported happiness by instrumenting for actual county income with county income predicted from national trends and county level occupation and industry composition. He finds very little difference between the OLS and IV results, suggesting such unobserved county shocks are not quantitatively significant. More generally, any story involving classical measurement error in our reference group income measures (relative to the unobserved true reference income) will imply attenuation bias (toward zero).

Finally, regarding the proportional hazards estimations, a common concern in such survival analysis is attenuation bias from unobserved individual heterogeneity. The concern is that individuals with especially negative individual effects ("frailty" in the parlance of survival analysis)–i.e., the θ_i term in our theoretical model–are more likely to exit the sample early via suicide; since there are no observations from these individuals for the remaining years of the sample, they receive less weight than survivors in the estimation, hence underestimating the effects of all variables on exit probability. Again, though, this bias only argues that the true effect of reference group income is in fact larger than what we find.

Our results confirm those obtained in semiaggregate analysis (Daly and Wilson 2006) on group suicide risk and income dispersion and also are broadly consistent with results using happiness surveys. The finding that reference income, holding own income constant, increases in suicide risk holds for reference groups ranging from simple geographic areas to near neighbors (evaluated as living in the same county and having one demographic marker in common) to simple geographical areas like county, with some evidence that age is particularly relevant for comparisons. State appears to be too broad as a measure of reference group. This finding is notable since many previous papers investigating relative income or relative deprivation have been forced to rely on state- or higher-level aggregates as reference groups (e.g., Blanchflower and Oswald 2004;

Kennedy, et al. 1996; and Kaplan, et al. 1996).

This paper has focused on static interpersonal income comparisons. Models of this kind are known by various names such as "external habit formation" and "Keeping Up with the Jones". Future research using suicide data may consider dynamic models of preferences such as "internal habit formation" or "Catching Up with the Jones". The evidence in this paper regarding the usefulness of suicide data for evaluating the nature of the utility function and preferences suggests that such research could indeed be fruitful.

References

Abel, Andrew B. “Asset Prices under Habit Formation and Catching up with the Joneses.” *American Economic Review*, May 1990, 80(2), pp. 38-42.

Alesina, Alberto, Rafael Di Tella, and Robert J. MacCulloch. "Inequality and Happiness: Are Europeans and Americans Different?" *Journal of Public Economics*, August 2004, v. 88, iss. 9-10, pp. 2009-42.

Alpizar, Francisco, Fredrik Carlsson, and Olof Johansson-Stenman. "How Much Do We Care About Absolute Versus Relative Income and Consumption?" *Journal of Economic Behavior & Organization*, 2005, 56, no. 3, pp. 405-21.

Becker, Gary S. and Luis Rayo. “Evolutionary Efficiency and Happiness.” *Journal of Political Economy*, April 2007, v. 115, iss. 2, pp. 302-37.

Bertrand, Marianne, and Sendhil Mullainathan. "Do People Mean What They Say? Implications for Subjective Survey Data." *American Economic Review*, 2001, 91(2), pp. 67-72.

Blakely, T. A., S. C. D. Collings, and J. Atkinson. "Unemployment and Suicide: Evidence for a Causal Association?" *Journal of Epidemiology and Community Health*, 2003, 57, pp. 594-600.

Blanchflower, David G., and Andrew J. Oswald. “Well-being over time in Britain and the USA.” *Journal of Public Economics*, 2004, 88, no. 7-8, pp. 1359-86.

Brainerd, Elizabeth. "Economic Reform and Mortality in the Former Soviet Union: A Study of the Suicide Epidemic in the 1990s." *European Economic Review*, 2001, 45, no. 4-6, pp. 1007-19.

Brekke, KA. (1997), *Economic Growth and the Environment: On the Measurement of Income and Welfare* (Cheltenham, Brookfield: Edward Elgar).

Brosnan, Sarah F., and Frans B. M. de Waal. "Monkeys Reject Unequal Pay." *Nature*, 2003, 425 297-99.

Brown, Gordon D.A., Jonathan Gardner, Andrew Oswald, and Jing Qian. "Does Wage Rank Affect Employees’ Well-being?" *IZA Discussion Papers,* March 2005, 1505, Institute for the Study of Labor (IZA).

Chuang, Hwei Lin, and Wei Chiao Huang. "Economic and Social Correlates of Regional Suicide Rates: A Pooled Cross-Section and Time-Series Analysis." *Journal of Socio Economics*, 1997, 26, no. 3 277-89.

Clark, Andrew E., and Andrew J. Oswald. "Satisfaction and Comparison Income." *Journal of Public Economics*, September 1996, 61(3), pp. 359-81.

Cutler, David, Edward Glaeser, and Karen Norberg. "Explaining the Rise in Teenage Suicide." *National Bureau of Economic Research Working Paper*, 2000, 7713.

Daly, Mary C., and Daniel J. Wilson. “Keeping Up with the Joneses and Staying Ahead of the Smiths: Evidence from Suicide Data.” Federal Reserve Bank of San Francisco Working Paper, 2006-12 (2006).

Diamond, Peter. Journal of Public Economics.

Di Tella, Rafael, Robert J. MacCulloch, and Andrew J. Oswald. "Preferences over Inflation and Unemployment: Evidence From Surveys of Happiness." *American Economic Review*, 2001, 91, no. 1 335-41.

Duesenberry, James S. *Income, Saving, and the Theory of Consumer Behavior*. Cambridge, MA: Harvard University Press, 1949.

Durkheim, Emile. 1951. Suicide: A Study in Sociology translated by George Simpson and John A. Spaulding. New York: The Free Press.

Easterlin, Richard A. "Does Money Buy Happiness?" *Public Interest*, 1973, XXX 3-10.

Easterlin, Richard A. "Does Economic Growth Improve the Human Lot? Some Empirical Evidence." *Nations and Households in Economic Growth: Essays in Honour of Moses Abramovitz.* Editors Paul A. David and M. W. RederNew York and London: Academic Press, 1974.

Easterlin, Richard A. "Will Raising the Incomes of All Increase the Happiness of All?" *Journal of Economic Behavior and Organization*, 1995, 27, pp. 35-48.

Engelmann, Dirk, and Martin Strobel. "Inequality Aversion, Efficiency, and Maximin Preferences in Simple Distribution Experiments." *American Economic Review*, September 2004, 94(4), pp. 857-69.

Ferrer-i-Carbonell, Ada. “Income and well-being: an empirical analysis of the comparison income effect.” *Journal of Public Economics*, June 2005, 89(5-6), pp. 997-1019.

Frank, Robert H. *Luxury Fever: Money and Happiness in an Era of Excess*. New York: The Free Press, 2000.

Gali, Jordi. “Keeping Up with the Joneses: Consumption Externalities, Portfolio Choice, and Asset Prices.” *Journal of Money, Credit, and Banking*, February 1994, 26(1) pp. 1-8.

Gerdtham, Ulf-G., and Magnus Johannesson. “Absolute Income, Relative Income, Income Inequality, and Mortality.” *Journal of Human Resources*, 2004, 39(1), pp. 228-47.

Hamermesh, Daniel S. “The Economics of Black Suicide.” *Southern Economic Journal*, 1974, 41, no. 2 188-99.

Hamermesh, Daniel S., and Neal M. Soss. "An Economic Theory of Suicide." *Journal of Political Economy*, January/February 1974, 82(1), pp. 83-98. (In *Mercurio* as "Una Teoria Economica del Suicidia.")

Helliwell, John F. "How's Life? Combining Individual and National Variables to Explain Subjective Well-Being." *Economic Modelling*, March 2003, v. 20, iss. 2, pp. 331-60.

Helliwell, John F. "Well-Being and Social Capital: Does Suicide Pose a Puzzle?" *National Bureau of Economic Research Working Paper*, 2004, no. 10896.

Hendin, Herbert. "Suicide in America." New York: W.W. Norton & Company, 1995.

Huang, Wei Chiao. "Religion, Culture, Economic and Sociological Correlates of Suicide Rates: A Cross-National Analysis." *Applied Economics Letters*, 1996, 3, no. 12 779-82.

Kahneman, Daniel and Amos Tversky. "On the Reality of Cognitive Illusions: A Reply to Gigerenzer's Critique." *Psychological Review*, 1996, 103, pp. 582-91.

Kaplan, George A., E.R. Pamuk, J.W. Lynch, Balfour Cohen, et al. "Inequality in Income and Mortality in the United States." *British Medical Journal*, 1996, 312, pp. 999-1003.

Kennedy, B.P., I. Kawachi, and D. Prothrow-Smith. "Income Distribution and Mortality: Cross Sectional Ecological Study of the Robin Hood Index in the United States." *British Medical Journal*, 1996, 312, pp. 1004-1007.

Kimenyi, Mwangi S., and William F. II Shughart. "Economics of Suicide: Rational or Irrational Choice." *Atlantic Economic Journal*, 1986, 14, no. 1 121.

Koivumaa-Honkanen, Heli, Risto Honkanen, Heimo Viinamaki, Kauko Heikkila, Jaakko Kaprio, and Markku Koskenvuo. "Life Satisfaction and Suicide: a 20-Year Follow-Up Study." *American Journal of Psychiatry*, 2001, 158(3), pp. 433-9.

Kposawa, Augustine J. "Unemployment and Suicide: A Cohort Analysis of Social Factors Predicting Suicide in the US National Longitudinal Mortality Study." *Psychological Medicine*, 2001, 31(1), pp. 127-38.

Lewis, Glyn, and Andy Sloggett. "Suicide, deprivation, and unemployment: record linkage study." *BMJ*, 1998, 317, pp. 1283-86.

Luttmer, Erzo F.P. "Neighbors as Negatives: Relative Earnings and Well-Being." *Quarterly Journal of Economics,* 2005, 102(3), pp. 963-1002.

Lykken, David. 1999 *Happiness: What Studies of Twins Show Us? Nature, Nurture, and the Happiness Set Point.* Golden Books Co. Inc.

Marcotte, Dave E. "The Economics of Suicide, Revisited." *Southern Economic Journal*, 2003, 69, no. 3, pp. 628-43.

Miller, Douglas, and Christina Paxson. "Relative Income, Race, and Mortality." *Journal of Health Economics*, 2006, 25(5), pp. 979-1003.

Miller, Mathew, Deborah Azrael, and David Hemenway. "Firearm Availability and Unintentional

Firearm Deaths, Suicide, and Homicide Among 5-14 Year Olds." *The Journal of Trauma*, 2002, 52(2), pp. 267-75.

Mohler, B., and F. Earls. "Trends in Adolescent Suicide: Misclassification Bias?" *American Journal of Public Health*, 2001, 91, no. 1 150-3.

Moyer, L. A., C. A. Boyle, and D. A. Pollock. "Validity of Death Certificates for Injury-Related Causes of Death." *American Journal of Epidemiology*, 1989, 130 1024-32.

Murphy, Kevin M., and Robert H. Topel. "Estimation and Inference in Two-Step Econometric Models." *Journal of Business and Economic Statistics*, 1985, 3, no. 4 370-79.

Osmani, S. (1993). 'Comment on B.M.S. van Praag: The relativity of the welfare concept', in (M. Nussbaum and A. Sen, eds.) *The Quality of Life*, Oxford: Oxford University Press.

Oswald, Andrew J. "Happiness and Economic Performance." *Economic Journal,* November 1997, 107(445), pp. 1815-31.

Peterson, C., N. Park, and M.E.P. Seligman. "Orientations to Happiness and Life Satisfaction: the Full Life Versus the Empty Life." *Journal of Happiness Studies*, 2005 6(1), 24-41.

Rockett, Ian, and G. S. Smith. "Suicide Misclassification in an International Context." How It's Done: An Invitation to Social Research. Editors B. S. Adler and R. D. Clark, 122-29. Belmont, CA: Wadsworth, 1999.

Ruggles, Stephen, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander. *Integrated Public Use Microdata Series: Version 3.0* [Machine-readable database]. Minneapolis, MN: Minnesota Population Center [producer and distributor], 2004. Available at http://www.ipums.org.

Ruhm, Christopher J. "Are Recessions Good For Your Health." *Quarterly Journal of Economics*, 2000, 115(2), pp. 617-50.

Schapiro, Morton Owen, and Dennis A. Ahlburg. "Suicide: The Ultimate Cost of Unemployment." *Journal of Post Keynesian Economics*, 1982-1983, 5(2) pp. 276-80.

Stevenson, Betsey, and Justin Wolfers. "'Til Death Do Us Part: Effects of Divorce Laws on Suicide, Domestic Violence and Spousal Murder." Mimeo, University of Pennsylvania, 2000.

Stevenson, Betsey, and Justin Wolfers. "Economic Growth and Subjective Well-Being: Reassessing the Easterlin Paradox." Mimeo, University of Pennsylvania, 2008.

Tekin, Erdal, and Sara Markowitz. "Suicidal Behavior and the Labor Market Productivity of Young Adults." NBER Working Paper #11238, 2005.

Tomes, Nigel. "Income Distribution, Happiness and Satisfaction: A Direct Test of the Interdependent Preferences Model." *Journal of Economic Psychology*, 1986, 7, no. 4 425-46.

Veblen, Thorsten (1899). *The Theory of the Leisure Class*, New York: Macmillan.

World Health Organization. WHO Mortality Database, January 2005. Available online: http://www3.who.int/whosis/menu.cfm?path=whosis,mort

U.S. Bureau of the Census. "National Longitudinal Mortality Study, Reference Manual, Version 2." August 1, 2005. Available online at: http://www.census.gov/nlms/reference.html.

U.S. Department of Health and Human Services. National Center for Health Statistics. Mortality Detail File, 1992 [Computer file]. Hyattsville, MD: U.S. Dept. of Health and Human Services, National Center for Health Statistics [producer], 1994. Ann Arbor, MI: Interuniversity Consortium for Political and Social Research [distributor], 1996.

Zimmerman, Frederick J., and Janice F. Bell. "Income Inequality and Physical and Mental Health: Testing Associations Consistent with Proposed Causal Pathways." *Journal of Epidemiology and Community Health*, 2006, 60, pp. 513-21.

Appendix A. Theoretical Model

In this Appendix, we describe in more detail the theoretical model that underlies our empirical analyses (which was discussed less formally in Section 3). The model embeds a dynamic programming model of the individual's suicide decision within the random utility model (RUM), the standard economic framework for modeling discrete choice.

The basic random utility model, $u(\mathbf{x}_{it}) + \theta_i$, lets contemporaneous utility for individual i in period t be a function of a deterministic component, $u()$, and an unobserved idiosyncratic component, θ_i, which can be treated as a random variable. The deterministic component $u()$ depends on various state variables contained in the vector $\mathbf{x}$. We assume $u()$ is monotonic in each variable in $\mathbf{x}$, strictly concave, and twice-differentiable. Mathematically, an individual's decision in each period whether or not to commit suicide can be expressed by the following dynamic programming problem:

$$V(\mathbf{x}_{it} \mid \theta_i) = \max_{S_{it}} \left\{ \left(1 - S_{it}\right) \left[u(\mathbf{x}_{it}) + \theta_i + \beta E_t V(\mathbf{x}_{i,t+1} \mid \theta_i) \right] \right\}$$

(A1)

subject to the boundary conditions

$$V(\mathbf{x}_{iT} \mid \theta_i) = 0;\ S_{it} = 1 \Rightarrow (S_{i,t+s} = 1)\ \forall\ s > 0 ,$$

and given a state transition process

$$\mathbf{x}_{it+1} = g(\mathbf{x}_{it}) + \varepsilon_{it+1} \text{ , where } \varepsilon_{it+1} \sim N(0, \sigma^2) .$$ [32]

In the above equations, S_{it} is the choice variable. It takes on the value 1 if suicide is chosen,

[32] In the present setup, there is no explicit role for age. To incorporate age effects, as in Hamermesh and Soss (1974), the Bellman equation could be augmented as follows to account for the probability of dying next period (by nonsuicide):

$$V(\mathbf{x}_{it} \mid \theta_i) = \max_{S_{it}} \left\{ \left(1 - S_{it}\right) \left[u(\mathbf{x}_{it}) + \theta_i + \beta P(T - t) E_t V(\mathbf{x}_{it+1} \mid \theta_i) \right] \right\},$$

where $P(T - t)$ is the probability of surviving to the next period given the individual's maximum remaining lifespan, $T - t$. It should be the case that $P(0) = 0$, $P(T\text{-}t)$. 0, and $P\,'(\,) > 0$. This implies that $V(\mathbf{x}_t)$ is decreasing in t, which is equivalent to age in this context. However, age could itself be a variable in $\mathbf{x}$ affecting utility. Therefore, our empirical analysis will allow for age effects with no restriction on sign.

0 otherwise. The objective function above assumes that the individual receives zero instantaneous utility if he or she chooses suicide. Setting the instantaneous utility from suicide equal to zero is an innocuous normalization; individual differences in this utility value can be thought of as part of θ_i, the individual-specific component of utility.[33] β is the discount factor and $E_tV(\mathbf{x}_{i,t+1}|\theta_i)$ is the expected value of future utility given this period's decision, S_t.

The first boundary condition states that the value function, V, is equal to zero in the final period, T, of the individual's natural life span. The second boundary condition states that suicide is an irreversible decision: choosing death today guarantees death in all future periods.

This dynamic programming problem can be solved recursively. Let $V_{it+1}^c(\mathbf{x}_{it+1} \mid \theta_i)$ denote the maximal value in period t+1 conditional on $S_{it} = 0$, and let $S_{i,t+1,t}$ denote the expectation as of t of the suicide decision next period conditional on $S_{it} = 0$, i.e., $S_{i,t+1,t} = E_t\,[S_{it+1}|\, S_{it}=0]$. This can be thought of as the individual's self-assessed probability of committing suicide next period. The solution for any period $t< T$ is:

$$S_{it}^* = \begin{Bmatrix} 1, \text{ if } u(\mathbf{x}_{it}) + \theta_i + \beta E_t V_{it+1}^c(\mathbf{x}_{it+1} \mid \theta_i) < 0 \\ 0, \text{ otherwise} \end{Bmatrix}, \text{ where} \tag{A2}$$

$$E_t V_{it+1}^c(\mathbf{x}_{it+1} \mid \theta_i) = E_t \left\{ \sum_{j=0}^{T-(t+1)} \beta^j \left[u(\mathbf{x}_{i,t+1+j}) + \theta_i \right] \left(1 - S_{i,t+1+j,t+j}\right) \right\}. \tag{A3}$$

This solution equation indicates that an individual will choose suicide when present and expected future utility (conditional on not choosing suicide this period) is less than the utility received from choosing suicide (recall that θ_i is the idiosyncratic utility component net of whatever instantaneous utility, either positive or negative, one receives upon committing suicide). The conditional expected

[33]The instantaneous utility from suicide may differ across individuals and could arguably be positive or negative (relative to natural death). Religious beliefs regarding the sinfulness of suicide or compassion for mourning loved ones left behind, for instance, would yield a negative instantaneous utility from suicide. On the other hand, one might receive some satisfaction from the sympathy and attention one might receive posthumously, implying a positive value.

future utility is the expected value of the discounted sum of future utilities, where each future utility value is weighted by the probability of not committing suicide in that period.

A key parameter in this model is θ_i, an unobserved, individual-specific component that incorporates all unobserved exogenous risk factors that determine an individual's predisposition to commit suicide. Conditional on $\mathbf{x}_{it}$ and β, whether $u(\mathbf{x}_{it}) + \theta_i + \beta E_t V^c_{it+1}(\mathbf{x}_{it+1} \mid \theta_i)$ is positive or negative will be completely determined by the value of the idiosyncratic component θ_i.[34] Lower levels of θ_i lower utility and increase the probability of suicide. In our framework, θ_i incorporates the various genetic, biochemical, psychological, and neurological preconditions that have been found to be prevalent in suicidal individuals. More generally, the individual-specific component of the randomized utility model allows for the fact that not all individuals will receive the same level of utility from the same level of **x**, even though **x** affects every individual in the same way. In other words, even if all individuals had the same levels of consumption, leisure, etc., some individuals would still be "happier" than others and some presumably would still commit suicide. The predisposition factors in θ_i are assumed to be pre-determined (genetically or formatively) for the adult age range on which we focus. Under this assumption, θ_i can be thought of as a random disturbance term that is predetermined and orthogonal to x. For a given $\mathbf{x}_{it}$, there exists a critical value $\overline{\theta}(\mathbf{x}_{it})$ below which $u(\mathbf{x}_{it}) + \theta_i + \beta E_t V^c_{it+1}(\mathbf{x}_{it+1} \mid \theta_i)$ is negative and the individual chooses to commit suicide. The effect of any variable contained in the $\mathbf{x}_{it}$ vector on the suicide decision can be found by partially differentiating $\overline{\theta}(\mathbf{x}_{it})$ with respect to that variable using the implicit function theorem[10]:

[34]Recall that from the state transition equation in (A1), $\mathbf{x}_{it+1}$ is a known stochastic function of $\mathbf{x}_{it}$, so $E_t V^c_{it+1}(\mathbf{x}_{it+1} \mid \theta_i)$ is known as of t.

[10] The implicit function is the solution for $\overline{\theta}$ of the equation:

$$\frac{\partial\bar{\theta}}{\partial x_{it}} = -\left[\underbrace{\frac{\partial u_t}{\partial x_{it}}}_{\text{A}} + \beta \underbrace{\frac{\partial E_t V_{t+1}^c(\mathbf{x}_{it+1})}{\partial x_{it+1}} \cdot \frac{\partial \mathbf{x}_{it+1}}{\partial x_{it}}}_{\text{B}}\right] / \frac{\partial F}{\partial\bar{\theta}} \quad . \tag{A4}$$

By construction, $\partial F / \partial\bar{\theta} > 0$, so the sign of $\partial\bar{\theta}/\partial x_{it}$ will equal the negative of the sign of the sum of the two expressions in the brackets of equation (4), which we've labeled **A** and **B**. **A** represents the effect of the variable x_{it} on contemporaneous utility; we call this the contemporaneous effect. **B** represents the effect that x_{it} has on future values of x_i as well as other variables in **x** (weighted by the marginal utility of x_i); we call this the signal effect. An example of the signal effect is that current income may serve as a signal of future income; thus, part of the negative effect of current income on suicide risk that we estimate in our forthcoming regressions may be due to the fact that current income is thought to be a good signal of future income and expectations of high future income may lower current suicide risk. For the variables considered in the empirical analysis of this paper, the signal effect is likely to be small or near zero and hence the overall effect will be dominated by the contemporaneous effect. Nonetheless, it is important to recognize that the effects of variables on utility and suicide that we identify represent the combined effect of these two forces.[11] The probability that an individual commits suicide is simply the probability that his/her θ_i is below the threshold level $\bar{\theta}(\mathbf{x}_{it})$. We describe how this probability is parameterized for our empirical analyses in Section 3.2.

$F(x_t, \bar{\theta}) = \left(u(\mathbf{x}_{it}) + \bar{\theta} + \beta E_t V^c(\mathbf{x}_{it+1} \mid \bar{\theta})\right) = 0$.

[11] The potential existence of the signal effect would not be recognized if one modeled individual behavior as depending on a single function representing remaining lifetime utility, as in Hamermesh and Soss (1974). In that case, the effect of a variable on lifetime utility is assumed to be unambiguous.

Appendix B. Construction of Geographic Aggregates

This appendix describes the construction of the geographic aggregate variables used in this study.

The county income data are based on the Census Bureau's Summary Table Files, SF-3, from the 1980, 1990, and 2000 decennial censuses. Note that income values reported in the 1980, 1990, and 2000 decennial censuses refer to income levels in 1979, 1989, and 1999, respectively. We measure county income for non-census years using the following interpolation procedure: (1) For each state and year, calculate the percentage deviation between that year's growth rate in Gross State Product (GSP) and the average (annualized) growth rate from T to T+10, for T = 1979, 1989; (2) Compute the average growth rate in county income from T to T+10; (3) Compute an estimated growth rate in county income as this 10-year average plus the percentage deviation from average in the county's state, as computed in step (2); (4) starting with county income in year T, compute county income in years T+1,...,T+9 using this estimated annual growth rate. This method preserves county differences in average growth over each decade but forces each county in a state to have parallel time series deviations from its decadal trend. Lastly, these nominal income levels were deflated to constant 1990 dollars using the CPI-U price index.

In some regressions below, we control for county-level cost of housing. Quality-adjusted house price indices are not available at the county level, so we constructed a hedonic house price index using data from the 1990 and 2000 PUMS data. The PUMS contains household-level data on house market value and numerous housing characteristics. The finest level of geographic detail in these data is the household's "Public-Use Microdata Area" (PUMA). Using the 1990 sample, we regressed log house value on PUMA fixed effects and a rich array of dummy variables covering all possible values of the housing characteristics variables, for all owner-occupied housing.[12] The

[12] The housing characteristics were property acreage, condo status, kitchen status, number of rooms, plumbing status, age of building, number of units in building, and number of bedrooms.

estimated PUMA fixed effects represent a constant-quality house price index for 1990. We used the estimated coefficients on the housing characteristics, each of which represents the percentage effect of the characteristic on house values, and the 2000 PUMS data on housing characteristics to obtain out-of-sample predicted house values for the 2000 PUMS observations. Averaging the difference between actual and predicted house value across households within PUMA yields a constant-quality house price index for 2000. The 2000 values are converted to 1990 dollars using the CPI-U. We use 1990 and 2000 PUMA-to-County mapping files from the Census Bureau to convert the real house price index from PUMA-level to County-level.[13] We obtain values for years 1979 to 1998 (the NLMS sample range) via linear interpolation and extrapolation from the 1990 and 2000 values. (Since the index represents the *logarithm* of real constant-quality housing values, linear interpolation amounts to assuming a constant within-county growth rate.)

Finally, we merge in data, from the Census Bureau's Summary Table Files, on shares of county population by race (white, black, other) and by broad age group (<20, 20-64, 65+).

[13] Counties that contain multiple pumas got the population-weighted average of those puma's index values; counties that shared a puma with other counties were all assigned that puma's fixed effect.

APPENDIX C

TABLE C1
Summary Statistics of Model Variables (NLMS)

Variable	Mean	S.D.	Max	Min
	A. Individual Income			
Own family income (in 1990 $)	40,775	34,619	2,013,388	-32,640
Income: Less than $10K	0.095	0.294	1	0
Income: $10K to $20K	0.149	0.356	1	0
Income: $20K to $40K	0.347	0.476	1	0
Income: $40K to $60K	0.225	0.418	1	0
(Omitted Category: More than $60K)				
	B. County income			
County income per family (in 1990 $)	48,025	12,306	164,625	14,467
	C. Individual Demographics			
Female	0.523	0.499	1	0
(Omitted Category: Male)				
Black	0.100	0.300	1	0
Other race	0.035	0.184	1	0
(Omitted Category: White)				
Age 20 - 24	0.137	0.344	1	0
Age 25 - 34	0.283	0.451	1	0
Age 35 - 44	0.235	0.424	1	0
Age 45 - 54	0.189	0.391	1	0
(Omitted Category: Age 55-64)				
Rural	0.315	0.465	1	0
(Omitted Category: Urban)				
	D. Individual Marital Status			
Widowed or Divorced	0.110	0.313	1	0
Separated	0.026	0.158	1	0
Single /never married	0.188	0.391	1	0
(Omitted Category: Married)				
	E. Individual Education			
Less than high school	0.175	0.380	1	0
More than high school	0.423	0.494	1	0
(Omitted Category: High School)				
	F. Individual Employment Status			
Employed but not working	0.034	0.182	1	0
Unemployed	0.046	0.210	1	0
Not in labor force, unable	0.017	0.129	1	0
Not in labor force, retired	0.214	0.410	1	0
(Omitted Category: Employed & working)				
	G. Individual Veteran Status			
Vietnam veterans	0.060	0.238	1	0
Korea veterans	0.033	0.178	1	0
WWII veterans	0.044	0.205	1	0
Other veterans	0.042	0.201	1	0
(Omitted Category: Nonveteran)				

TABLE C2

Summary Statistics of Model Variables (MDF-PUMS)

Variable	Mean	S.D.	Max	Min
	A. Income			
Own Income (estimated)	47,793	21,250	494,060	5
County Income (mean family)	46,465	9,769	79,592	20,859
County*Age Income (mean family)	46,607	11,999	90,687	9,436
County*Race Income (mean family)	46,340	11,076	93,385	21
	B. Individual Demographics			
Female	0.514	0.500	1	0
(Omitted Category: Male)				
Other Race	0.049	0.208	1	0
Black	0.119	0.313	1	0
(Omitted Category: White)				
Age 20-24	0.118	0.330	1	0
Age 25-34	0.289	0.456	1	0
Age 35-44	0.263	0.439	1	0
Age 45-54	0.178	0.380	1	0
(Omitted Category: Age 55 - 64)	0.152	0.353	1	0
	C. Individual Marital Status			
Divorced	0.128	0.334	1	0
Widowed	0.024	0.150	1	0
Single /never married	0.229	0.423	1	0
(Omitted Category: Married)				
	D. Individual Education			
< 9th grade ed	0.035	0.234	1	0
Some HS	0.078	0.279	1	0
Some college	0.306	0.457	1	0
College degree	0.167	0.363	1	0
MA, prof degree, or PhD	0.085	0.271	1	0
(Omitted Category: 12th grade or GED)				
	E. Cell-Level Controls[a]			
Unemployment rate	0.053	0.084	1	0
Share of HH's living in own home	0.682	0.243	1	0
Share that are vietnam veterans	0.060	0.115	1	0
	F. State-Level Control			
Share of suicides committed by firearm	0.549	0.120	0.800	0.290

[a]Cell is defined by county-age-race-sex-Hispanic-education-marital status

FIGURE 1.
Schematic Relating Suicide Choice to Placement on Utility Spectrum

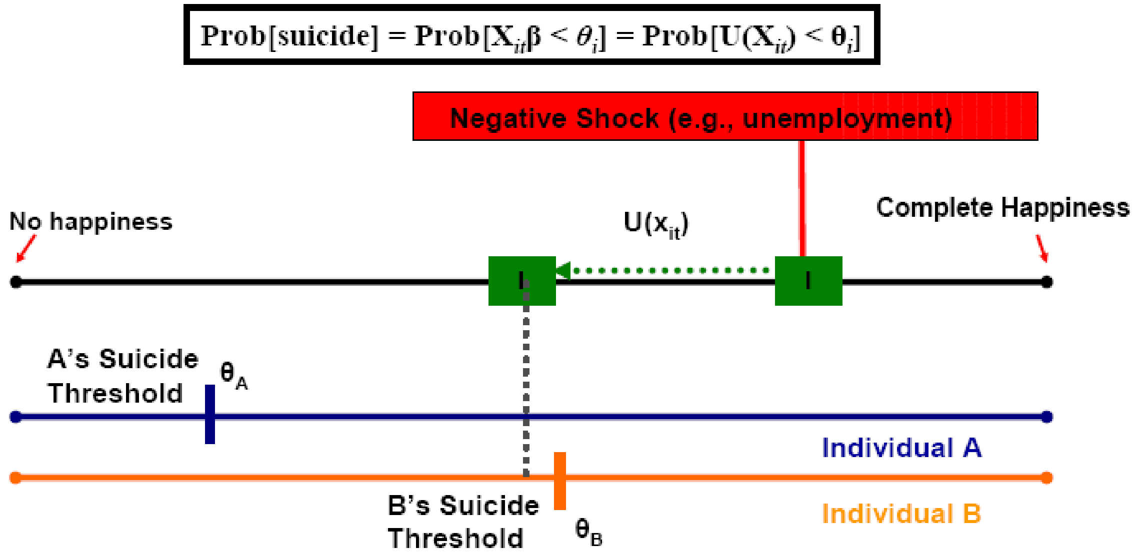

FIGURE 2
Actual Income Distribution for General and Suicide Populations, NLMS Sample

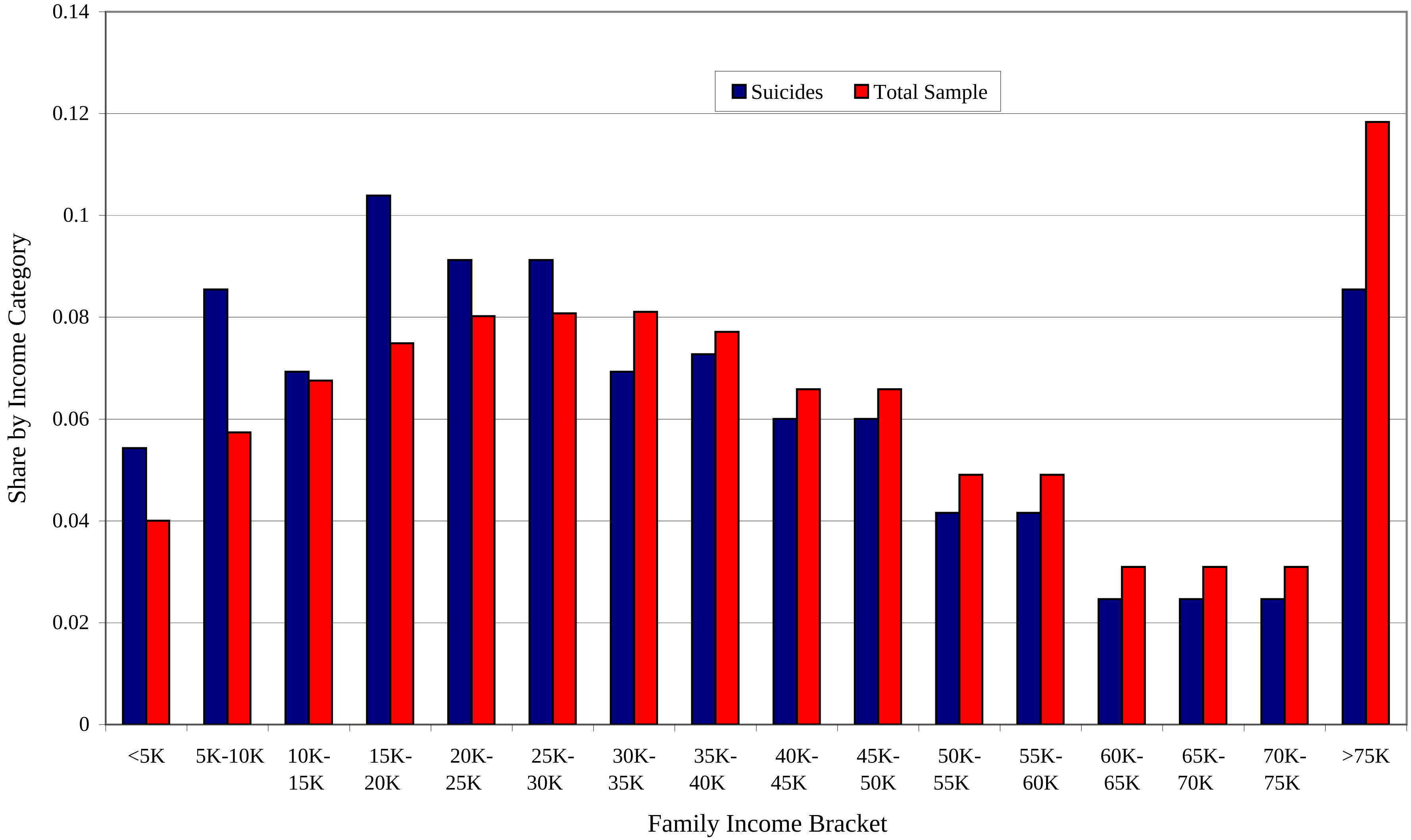

FIGURE 3
Estimated Income Distribution for General and Suicide Populations, MDF-PUMS Sample
(Kernel Density Estimates)

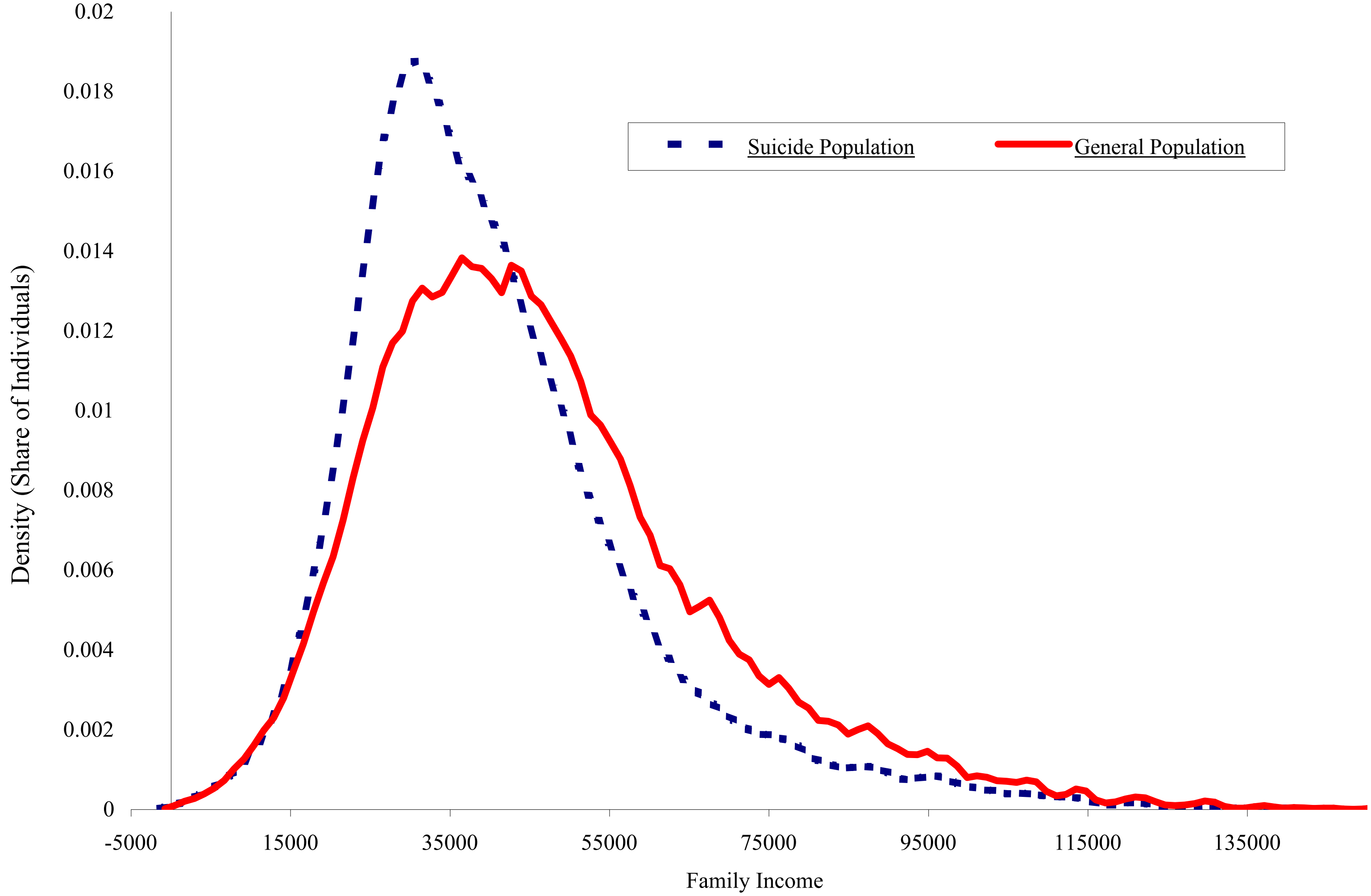

Figure 4. County Suicide Rates vs. County Income in 1990

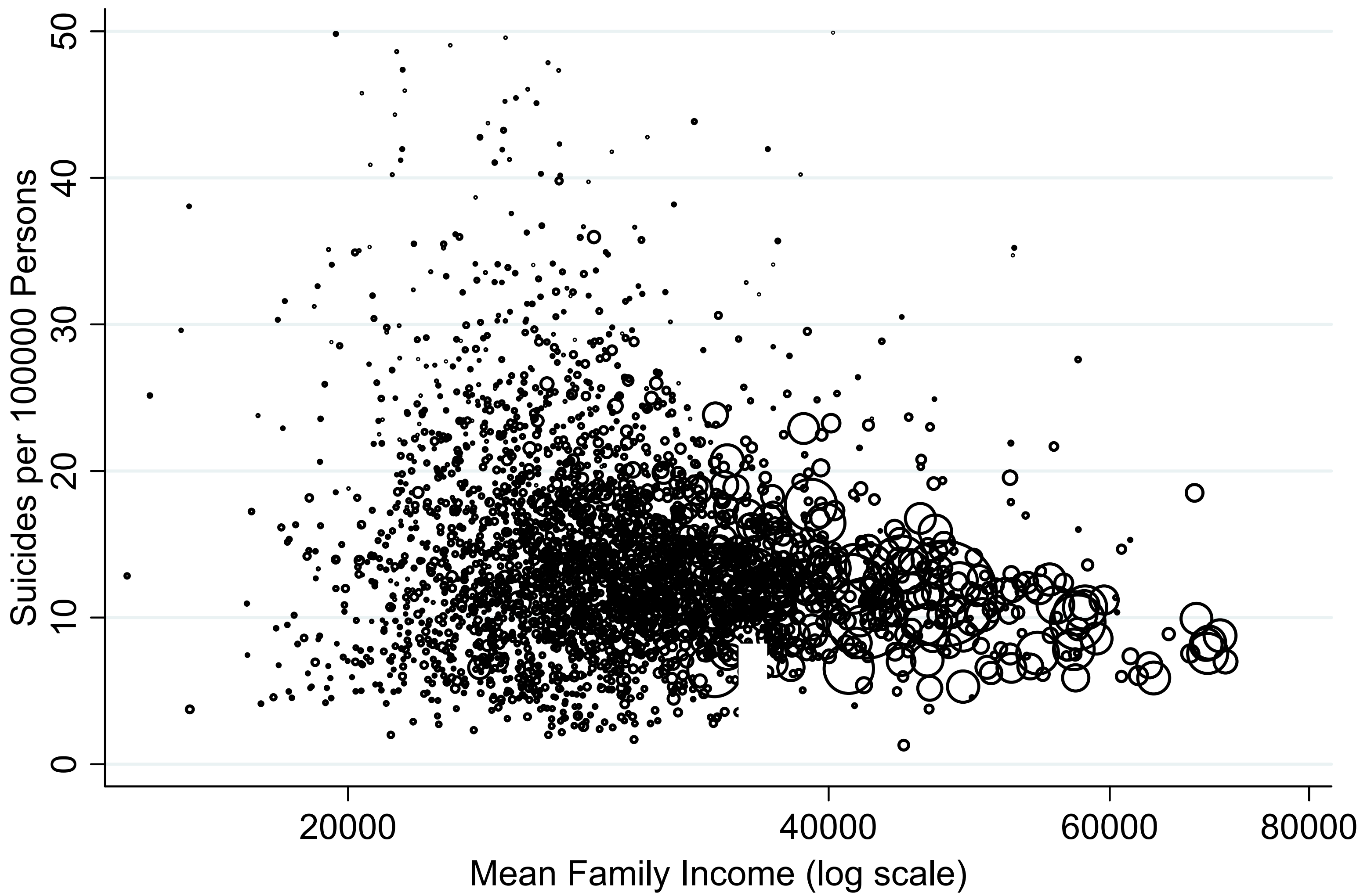

TABLE 1
Sample Suicide Rates, by Category
(Average Annual Rate per 100,000)

<table>
<tr><th></th><th colspan="2">Data Sets</th></tr>
<tr><th></th><th>NLMS[a]</th><th>MDF-PUMS</th></tr>
<tr><td></td><td colspan="2">A. Total Sample</td></tr>
<tr><td></td><td>12.98</td><td>13.37</td></tr>
<tr><td></td><td colspan="2">B. Age</td></tr>
<tr><td>Age: 20-24</td><td>12.79</td><td>13.80</td></tr>
<tr><td>Age: 25-34</td><td>13.14</td><td>13.53</td></tr>
<tr><td>Age: 35-44</td><td>12.91</td><td>13.52</td></tr>
<tr><td>Age: 45-54</td><td>12.21</td><td>13.22</td></tr>
<tr><td>Age: 55-64</td><td>14.00</td><td>12.66</td></tr>
<tr><td></td><td colspan="2">C. Race</td></tr>
<tr><td>Race: White</td><td>14.05</td><td>14.37</td></tr>
<tr><td>Race: Black</td><td>6.51</td><td>9.38</td></tr>
<tr><td>Race: Other</td><td>7.06</td><td>6.21</td></tr>
<tr><td></td><td colspan="2">D. Gender</td></tr>
<tr><td>Sex: Male</td><td>21.06</td><td>21.46</td></tr>
<tr><td>Sex: Female</td><td>5.61</td><td>5.73</td></tr>
<tr><td></td><td colspan="2">E. Education</td></tr>
<tr><td>Educ: < 9th grade</td><td rowspan="2">16.83</td><td>13.89</td></tr>
<tr><td>Educ: Some HS</td><td>22.24</td></tr>
<tr><td>Educ: 12th grade or GED</td><td>13.11</td><td>18.44</td></tr>
<tr><td>Educ: Some college</td><td rowspan="3">11.20</td><td>9.00</td></tr>
<tr><td>Educ: College degree</td><td>8.50</td></tr>
<tr><td>Educ: MA, prof. degree, or PhD</td><td>10.80</td></tr>
<tr><td></td><td colspan="2">F. Marital Status</td></tr>
<tr><td>Married</td><td>11.72</td><td>8.90</td></tr>
<tr><td>Divorced</td><td>18.03</td><td rowspan="2">23.21</td></tr>
<tr><td>Separated</td><td>14.22</td></tr>
<tr><td>Widowed</td><td>10.01</td><td>16.44</td></tr>
<tr><td>Single/Never Married</td><td>15.23</td><td>19.66</td></tr>
</table>

[a]NLMS sample suicide rates are number of suicides during the follow-up period divided by total number of observations, divided by average exposure time (12.7 yrs).

TABLE 2

Baseline Models, NLMS Sample

	Dependent Variable: Suicide Death							
	Linear Own Income		Nonlinear Own Income		Own and Others' Income		Own and Others' Income + County Demographics	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
A. Income								
Log own family income	-0.080 ***	0.001						
Own family income: <$10K			0.367 ***	0.001	0.407 ***	0.001	0.419 ***	0.001
Own family income: $10K to $20K			0.338 ***	0.001	0.374 ***	0.000	0.386 ***	0.000
Own family income: $20K to $40K			0.076	0.375	0.102	0.254	0.110	0.219
Own family income: $40K to $60K			-0.025	0.786	0.012	0.897	-0.007	0.945
(Omitted Category: >$60K)								
Log county income per family					0.308 **	0.023	0.392 ***	0.009
B. County Controls								
County population share, Age 20-39							-0.409	0.653
County population share, Age 40-59							1.662 **	0.032
County population share, Age 60+							-0.516	0.321
(Omitted Category: share, Age <20)								
County population share, Black							0.415 **	0.010
County population share, Other Race							-0.042	0.842
(Omitted Category: share, White)								
C. Individual Controls								
Female	-1.395 ***	0.000	-1.400 ***	0.000	-1.400 ***	0.000	-1.403 ***	0.000
(Omitted Category: Male)								
Black	-0.774 ***	0.000	-0.789 ***	0.000	-0.787 ***	0.000	-0.854 ***	0.000
Other Race	-0.576 ***	0.004	-0.511 ***	0.006	-0.526 ***	0.002	-0.530 ***	0.004
(Omitted Category: White)								
Age 20 - 24	-0.221 *	0.076	-0.207 *	0.095	-0.193	0.120	-0.200	0.105
Age 25 - 34	-0.050	0.629	-0.047	0.654	-0.042	0.680	-0.049	0.620
Age 35 - 44	0.115	0.264	0.127	0.221	0.129	0.200	0.125	0.214
Age 45 - 54	-0.012	0.908	0.002	0.986	0.005	0.956	0.003	0.981
(Omitted Category: Age 55-64)								
Rural	-0.038	0.529	-0.062	0.287	-0.009	0.891	0.017	0.805
(Omitted Category: Urban)								
Widowed or divorced	0.591 ***	0.000	0.570 ***	0.000	0.558 ***	0.000	0.558 ***	0.000
Separated	0.507 ***	0.002	0.472 ***	0.003	0.460 ***	0.008	0.451 ***	0.009
Single /never married	0.370 ***	0.000	0.317 ***	0.000	0.301 ***	0.000	0.297 ***	0.000
(Omitted Category: Married)								
Less than high school	0.126 *	0.089	0.087	0.240	0.098	0.177	0.092	0.203
More than high school	-0.127 **	0.042	-0.094	0.134	-0.103	0.110	-0.107 *	0.098
(Omitted Category: high school)								
Employed but not working	0.303 **	0.020	0.284 **	0.030	0.286 **	0.029	0.286 **	0.028
Unemployed	0.577 ***	0.000	0.523 ***	0.000	0.524 ***	0.000	0.529 ***	0.000
Not in labor force, unable	0.880 ***	0.000	0.796 ***	0.000	0.798 ***	0.000	0.802 ***	0.000
Not in labor force, retired	0.415 ***	0.000	0.387 ***	0.000	0.390 ***	0.000	0.388 ***	0.000
(Omitted Category: Employed & working)								
Vietnam veterans	0.279 ***	0.002	0.309 ***	0.001	0.312 ***	0.001	0.308 ***	0.001
Korea veterans	0.255 **	0.042	0.265 **	0.035	0.269 **	0.032	0.264 **	0.035
WWII veterans	0.208 *	0.081	0.213 *	0.073	0.223 *	0.059	0.211 *	0.074
Other veterans	0.078	0.495	0.094	0.415	0.094	0.393	0.093	0.399
(Omitted Category: Nonveteran)								
Number of Suicides	1,401		1,401		1,401		1,401	
Number of Observations	957,796		957,796		957,796		957,796	

Notes: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. P-values are based on standard errors that are robust to heteroskedasticity and clustering within county.

TABLE 3
Robustness Checks on Baseline Models, NLMS Sample
Dependent Variable: Suicide Death

	Endogenous Own Income			
	Baseline Actual Own Income		Two-Stage Predicted Own Income	
	Coefficient	p-value	Coefficient	p-value
Log own family income	-0.088 ***	0.000	-0.632 ***	0.007
Log county income per family	0.333 **	0.025	0.458 ***	0.005
Number of Suicides	1,401		1,401	
Number of Observations	957,796		957,796	

Notes: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. P-values are based on standard errors that are robust to heteroskedasticity and clustering within county. All regressions in this table include the same set of individual and county control variables shown in 4th (rightmost) column of Table 2. Coefficients on control variables are not shown (available upon request).

TABLE 4
Robustness Checks on Baseline Models, NLMS Sample
Dependent Variable: Suicide Death

	Own and Others' Income + County Demographics			Cost of Living			Cost of Living			Cost of Living		
				State Fixed Effects			House Price Index			Are Renters Different?		
	Coefficient		p-value	Coefficient		p-value	Coefficient		p-value	Coefficient		p-value
Own family income: <$10K	0.419	***	0.001	0.400	***	0.001	0.415	***	0.001	0.390	***	0.002
Own family income: $10K to $20K	0.386	***	0.000	0.364	***	0.000	0.385	***	0.000	0.366	***	0.001
Own family income: $20K to $40K	0.110		0.219	0.100		0.257	0.111		0.216	0.102		0.272
Own family income: $40K to $60K	-0.007		0.945	-0.010		0.917	-0.004		0.964	-0.006		0.949
Log county income per family	0.392	***	0.009	0.437	***	0.007	0.565	***	0.001	0.428	**	0.014
County House Price Index							-0.118	*	0.070			
(Log county income per family)*(Renter)										-0.133		0.601
Renter										1.491		0.583
Number of Suicides	1,401			1,401			1,401			1,401		
Number of Observations	957,796			957,796			957,796			957,796		

Notes: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. P-values are based on standard errors that are robust to heteroskedasticity and clustering within county. All regressions in this table include the same set of individual and county control variables shown in the 4th (rightmost) column of Table 2. Coefficients on control variables are not shown (available upon request).

TABLE 5
Alternative Causes of Death

	Dependent Variable:							
	Suicide Baseline Model		Suicide + Undetermined Baseline Model		Heart Attack Baseline Model		All Cause Baseline Model	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
	A. Income							
Own family income: <$10K	0.419 ***	0.001	0.396 ***	0.000	0.441 ***	0.000	0.418 ***	0.000
Own family income: $10K to $20K	0.386 ***	0.000	0.378 ***	0.000	0.388 ***	0.000	0.345 ***	0.000
Own family income: $20K to $40K	0.110	0.219	0.102	0.218	0.248 ***	0.000	0.202 ***	0.000
Own family income: $40K to $60K	-0.007	0.945	-0.008	0.928	0.106 **	0.012	0.064 ***	0.000
Log county income per family	0.392 ***	0.009	0.372 ***	0.008	-0.548 ***	0.000	-0.026	0.334
	B. County Controls							
County population share, Age 20-39	-0.409	0.653	0.283	0.749	-1.945 ***	0.000	-0.022	0.899
County population share, Age 40-59	-1.662 **	0.032	-0.878	0.189	-0.185	0.555	-0.045	0.674
County population share, Age 60+	-0.516	0.321	-0.056	0.913	-0.567 **	0.014	-0.075	0.444
County population share, Black	0.415 **	0.010	0.482 ***	0.002	0.054	0.429	0.183 ***	0.000
County population share, Other Race	-0.042	0.842	-0.089	0.628	-0.069	0.481	0.078 **	0.014
	C. Individual Controls							
Female	-1.403 ***	0.000	-1.380 ***	0.000	-1.019 ***	0.000	-0.647 ***	0.000
(Omitted Category: Male)								
Black	-0.854 ***	0.000	-0.669 ***	0.000	-0.161 ***	0.001	0.185 ***	0.000
Other Race	-0.530 ***	0.004	-0.460 ***	0.009	-0.366 ***	0.000	-0.176 ***	0.000
(Omitted Category: White)								
Age 20 - 24	-0.200	0.105	-0.097	0.400	-4.819 ***	0.000	-3.019 ***	0.000
Age 25 - 34	-0.049	0.620	0.050	0.608	-3.251 ***	0.000	-2.452 ***	0.000
Age 35 - 44	0.125	0.214	0.156	0.111	-1.753 ***	0.000	-1.556 ***	0.000
Age 45 - 54	0.003	0.981	0.022	0.817	-0.742 ***	0.000	-0.682 ***	0.000
(Omitted Category: Age 55-64)								
Rural	0.017	0.805	-0.025	0.707	-0.094 ***	0.001	-0.093 ***	0.000
(Omitted Category: Urban)								
Widowed or Divorced	0.558 ***	0.000	0.566 ***	0.000	0.257 ***	0.000	0.301 ***	0.000
Separated	0.451 ***	0.009	0.507 ***	0.001	0.057	0.457	0.265 ***	0.000
Single /never married	0.297 ***	0.000	0.354 ***	0.000	0.050	0.326	0.275 ***	0.000
(Omitted Category: Married)								
Less than high school	0.092	0.203	0.179 ***	0.010	0.235 ***	0.000	0.175 ***	0.000
More than high school	-0.107 *	0.098	-0.133 **	0.040	-0.280 ***	0.000	-0.215 ***	0.000
(Omitted Category: high school)								
Employed but not working	0.286 **	0.028	0.314 **	0.013	0.161 ***	0.002	0.190 ***	0.000
Unemployed	0.529 ***	0.000	0.540 ***	0.000	0.192 ***	0.002	0.279 ***	0.000
Not in labor force, unable	0.802 ***	0.000	0.835 ***	0.000	1.018 ***	0.000	1.068 ***	0.000
Not in Labor Force, Retired	0.388 ***	0.000	0.417 ***	0.000	0.485 ***	0.000	0.476 ***	0.000
(Omitted Category: Employed & working)								
Vietnam veterans	0.308 ***	0.001	0.349 ***	0.000	-0.022	0.698	0.044 **	0.026
Korea veterans	0.264 **	0.035	0.270 **	0.026	0.031	0.482	0.086 ***	0.000
WWII veterans	0.211 *	0.074	0.209 *	0.071	0.172 ***	0.000	0.169 ***	0.000
Other veterans	0.093	0.399	0.133	0.211	-0.004	0.931	0.016	0.445
(Omitted Category: Nonveteran)								
Number of Deaths	1,401		1,546		8,603		74,786	
Number of Observations	957,796		957,796		957,796		957,796	

Notes: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. P-values are based on standard errors that are robust to heteroskedasticity and clustering within county.

TABLE 6
Extensions to Baseline Model, NLMS Sample
Dependent Variable: Suicide Death

	Baseline County Income		Alternative Reference State Income		Heterogeneous Effects across Income Distribution	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
			A. Income			
Own family income: <$10K	0.419 ***	0.001	0.383 ***	0.002	-1.749	0.710
Own family income: $10K to $20K	0.386 ***	0.000	0.354 ***	0.000	-2.983	0.457
Own family income: $20K to $40K	0.110	0.219	0.088	0.325	-1.195	0.758
Own family income: $40K to $60K	-0.007	0.945	-0.017	0.859	0.516	0.899
Log county income per family	0.392 ***	0.009			0.272	0.365
Log state income per family			0.080	0.693		
			B. Interactions			
(Log county income per family)*(Own Income: <$10K)					0.202	0.646
(Log county income per family)*(Own Income: $10K to $20K)					0.314	0.399
(Log county income per family)*(Own Income: $20K to $40K)					0.121	0.738
(Log county income per family)*(Own Income: $40K to $60K)					-0.049	0.895
Number of Suicides	1,401		1,401		1,401	
Number of Observations	957,796		957,796		957,796	

Notes: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. P-values are based on standard errors that are robust to heteroskedasticity and clustering within county. All regressions in this table include the same set of individual and county control variables shown in 4th (rightmost) column of Table 2. Coefficients on control variables not shown (available upon request).

TABLE 7

Baseline Suicide Risk and Reference Income, MDF-PUMS Sample

	Linear Own Income			Nonlinear Own Income			Own and Others' Income			State Fixed Effects			House Price Index		
	Marginal effect		p-value	Marginal effect		p-value	Marginal effect		p-value	Marginal effect		p-value	Marginal effect		p-value
A. Income															
Log Own family income	-1.57E-04	***	0.000												
Own family income: < 20th percentile				3.25E-04	***	0.000	4.04E-04	***	0.000	3.41E-04	***	0.000	4.15E-04	***	0.000
Own family income: 20th to 40th percentile				3.27E-04	***	0.000	4.11E-04	***	0.000	3.43E-04	***	0.000	4.22E-04	***	0.000
Own family income: 40th to 60th percentile				2.01E-04	***	0.000	2.57E-04	***	0.000	2.03E-04	***	0.000	2.63E-04	***	0.000
Own family income: 60th to 80th percentile				9.37E-05	***	0.000	1.20E-04	***	0.000	9.25E-05	***	0.000	1.23E-04	***	0.000
(Omitted Category: >80th percentile)															
Log County income per family							2.45E-04	***	0.000	2.34E-04	***	0.000	1.90E-04	***	0.007
County House Price Index													4.12E-05		0.244
B. Individual Controls															
Female	-2.59E-04	***	0.000	-2.46E-04	***	0.000	-2.42E-04	***	0.000	-2.05E-04	***	0.000	-2.36E-04	***	0.000
Race: Other	-2.17E-04	***	0.000	-1.96E-04	***	0.000	-1.95E-04	***	0.000	-1.74E-04	***	0.000	-1.93E-04	***	0.000
Race: Black	-2.20E-04	***	0.000	-2.05E-04	***	0.000	-2.04E-04	***	0.000	-1.62E-04	***	0.000	-1.99E-04	***	0.000
Age 20 - 24	-1.22E-04	***	0.000	-1.10E-04	***	0.000	-1.00E-04	***	0.000	-8.35E-05	***	0.000	-8.99E-05	***	0.000
Age 25 - 34	-7.79E-05	***	0.000	-7.59E-05	***	0.000	-7.07E-05	***	0.000	-5.89E-05	***	0.000	-6.17E-05	***	0.000
Age 35 - 44	5.19E-06		0.657	1.65E-05		0.140	1.67E-05		0.120	2.14E-05	***	0.001	2.24E-05	***	0.007
Age 45 - 54	4.07E-05	***	0.000	6.04E-05	***	0.000	6.52E-05	***	0.000	5.97E-05	***	0.000	6.82E-05	***	0.000
C. Cell Controls[a]															
Unemployment rate in cell	5.51E-04	***	0.000	5.41E-04	***	0.000	5.41E-04	***	0.000	4.65E-04	***	0.000	5.34E-04	***	0.000
Share of cell households that live in own home	-3.02E-04	***	0.000	-2.17E-04	***	0.002	-1.71E-04	***	0.001	-1.26E-04	***	0.000	-1.26E-04	***	0.009
Share of cell that are Vietnam veterans	-4.85E-05		0.136	-5.12E-05		0.116	-2.26E-05		0.129	-5.12E-05	***	0.001	-2.50E-05		0.272
D. State Control															
Share of suicides committed by firearm	4.44E-04	***	0.000	3.66E-04	***	0.000	5.03E-04	***	0.000	---	***	0.000	5.42E-04	***	0.000
Estimated Prob(suicide)	3.52E-04			3.34E-04			3.28E-04			3.34E-04			3.28E-04		
Number of observations	4,360,731			4,360,731			4,360,731			4,360,731			4,360,731		

Notes: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. P-values are based on standard errors that are robust to heteroskedasticity and allow for clustering within county.

[a]Cell is defined by county-age-race-sex-hispanic-education-marital status

TABLE 8
Alternative Reference Groups, MDF-PUMS Sample

	County		County*Age		County*Race		State	
	Marginal effect	p-value	Marginal effect	p-value	Marginal effect	p-value	Marginal effect	p-value
				A. Income				
Own family income: < 20th percentile	4.04E-04 ***	0.000	4.32E-04 ***	0.000	3.70E-04 ***	0.000	3.54E-04 ***	0.000
Own family income: 20th to 40th percentile	4.11E-04 ***	0.000	4.36E-04 ***	0.000	3.74E-04 ***	0.000	3.55E-04 ***	0.000
Own family income: 40th to 60th percentile	2.57E-04 ***	0.000	2.72E-04 ***	0.000	2.33E-04 ***	0.000	2.19E-04 ***	0.000
Own family income: 60th to 80th percentile	1.20E-04 ***	0.000	1.26E-04 ***	0.000	1.09E-04 ***	0.000	1.02E-04 ***	0.000
(Omitted Category: >80th percentile)								
(log) Mean family income, county	2.45E-04 ***	0.000						
(log) Mean family income, county*age			3.05E-04 ***	0.000				
(log) Mean family income, county*race					2.10E-04 ***	0.000		
(log) Mean family income, state							5.39E-05	0.280
				B. Individual Controls				
Female	-2.42E-04 ***	0.000	-2.42E-04 ***	0.000	-2.20E-04 ***	0.000	-1.55E-04 ***	0.000
Race: Other	-1.95E-04 ***	0.000	-1.96E-04 ***	0.000	-1.69E-04 ***	0.000	-7.98E-05 ***	0.000
Race: Black	-2.04E-04 ***	0.000	-2.06E-04 ***	0.000	-1.47E-04 ***	0.000	-8.94E-05 ***	0.000
Age 20 - 24	-1.00E-04 ***	0.000	-3.57E-05	0.143	-9.00E-05 ***	0.000	-3.88E-05 ***	0.000
Age 25 - 34	-7.07E-05 ***	0.000	-3.24E-05 *	0.072	-6.28E-05 ***	0.000	-2.42E-05 ***	0.000
Age 35 - 44	1.67E-05	0.120	-3.78E-06	0.718	1.59E-05	0.107	3.22E-06	0.348
Age 45 - 54	6.52E-05 ***	0.000	7.52E-06	0.481	5.98E-05 ***	0.000	1.39E-05 ***	0.000
				C. Cell Controls[a]				
Unemployment rate in cell	5.51E-04 ***	0.000	5.23E-04 ***	0.000	4.95E-04 ***	0.000	-8.62E-05 ***	0.000
Share of cell households that live in own home	-3.02E-04 ***	0.000	-1.72E-04 ***	0.007	-1.50E-04 **	0.017	1.74E-04 ***	0.000
Share of cell that are Vietnam veterans	-4.85E-05	0.136	-3.33E-05	0.279	-1.79E-05	0.533	-1.80E-05 *	0.065
				D. State Control				
Share of suicides committed by firearm	4.44E-04 ***	0.000	5.38E-04 ***	0.000	4.27E-04 ***	0.000	1.89E-04 ***	0.000
Estimated Prob(suicide)	3.52E-04		1.03E-04		1.03E-04		1.03E-04	
Number of observations	4,360,731		4,360,747		4,360,747		4,360,747	

Notes: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. P-values are based on standard errors that are robust to heteroskedasticity and allow for clustering within county.
[a]Cell is defined by county-age-race-sex-hispanic-education-marital status