FEDERAL RESERVE BANK OF SAN FRANCISCO

WORKING PAPER SERIES

Keeping Up with the Joneses and
Staying Ahead of the Smiths:
Evidence from Suicide Data

Mary Daly
Federal Reserve Bank of San Francisco

and

Dan Wilson
Federal Reserve Bank of San Francisco

May 2006

Keeping Up with the Joneses and Staying Ahead of the Smiths
Evidence from Suicide Data

Mary C. Daly
Economic Research Department
Federal Reserve Bank of San Francisco
101 Market Street
San Francisco, CA 94105
(415) 974-3186
mary.daly@sf.frb.org

Daniel J. Wilson
Economic Research Department
Federal Reserve Bank of San Francisco
101 Market Street
San Francisco, CA 94105
(415) 974-3423
daniel.wilson@sf.frb.org

Fax: 415-974-2168

First Draft: June 2005
Current Draft: May 2006

Keeping Up with the Joneses and Staying Ahead of the Smiths
Evidence from Suicide Data

Abstract

This paper empirically assesses the theory of interpersonal income comparison using a unique data set on suicide deaths in the United States. We treat suicide as a choice variable, conditional on exogenous risk factors, reflecting one's assessment of current and expected future utility. Using this framework we examine whether differences in group-specific suicide rates are systematically related to income dispersion, controlling for socio-demographic characteristics and income level. The results strongly support the notion that individuals consider relative income in addition to absolute income when evaluating their own utility. Importantly, the findings suggest that relative income affects utility in a two-sided manner, meaning that individuals care about the incomes of those above them (the Joneses) and those below them (the Smiths). Our results complement and extend those from studies using subjective survey data or data from controlled experiments.

Keeping Up with the Joneses and Staying Ahead of the Smiths
Evidence from Suicide Data

1. Introduction

A large and growing theoretical literature, spanning several disciplines, posits that individuals'

happiness is affected by their relative standing–the relative income hypothesis.[1] Empirical work on the

importance of relative income is much more limited and generally reliant on analyses of subjective well-

being surveys or observations from controlled experiments. While these studies point to a role for

relative income in the utility function, their subjective nature (in the case of happiness surveys) or

artificial and nonrandom environments (in the case of experiments) limit their acceptance as evidence of

the relative income hypothesis. In the absence of strong proof, most theoretical and empirical researchers

continue to model individuals as considering only their own income in formulating utility.

Making such an assumption is not costless, however. Several authors have shown that including

relative income in standard models of individual utility can alter conventional wisdom regarding optimal

economic interventions.[2] Others have argued that puzzling macro wage dynamics and unemployment

fluctuations (Akerlof and Yellen (1990) and movements in asset prices (Abel 1990) can be explained by

allowing for relative income effects on individual utility. Still others have noted that whether economic

growth is welfare enhancing or not depends on whether individuals are allowed to compare themselves to

a benchmark level of consumption (Easterlin 1974). Finally, much of the debate over rapid increases in

income inequality hinges on whether individuals care about their relative ranking. If preferences are

interdependent, increases in inequality can have an independent negative effect on welfare even when

absolute income rises for everyone.

---

[1]For example Veblen 1899; Duesenberry 1952; Easterlin 1974; Abel 1990; Gali 1994; Kahneman and Tversky 1996; Becker and Rayo (2004).

[2]For example see Boskin and Sheshinski 1978; Layard 1980; Oswald 1984; Ng 1987; Seidman 1987; Ireland 1998; Ljungqvist and Uhlig 2000; and Abel 2003.

The importance of this unsettled issue suggests that further examinations are warranted. We contribute to this literature in two ways. Responding to limitations inherent in survey and experimental data, we first examine an alternative measure of individual happiness (unhappiness)–suicide deaths. Treating suicide as a "revealed preference" regarding current life satisfaction and assessed value of future life, we examine how utility is affected by own and relative income. We then extend the typical empirical test of the relative income hypothesis, which focuses on overall inequality, and investigate whether relative income effects differ across the income distribution; specifically, we consider whether inequality in the upper and lower tails affects suicide risk differently.

Our basic model relates group-level suicide rates, where groups are defined by county×age×race×sex cells, to group-level controls and various county-level measures of absolute and relative income. Our use of county-level income measures follows others in the literature and assumes that the relevant reference group for interpersonal comparison is a local geographic area. Consistent with prior work, results from this model confirm a weak role for overall inequality. The importance of income dispersion reveals itself when overall inequality is decomposed into its upper and lower tail components. In these regressions, inequality has a statistically significant and economically important effect on suicide risk. We find that increases in upper tail inequality raise suicide risk while increases in lower tail inequality lower suicide risk. These results hold across an array of robustness checks on econometric specification, data and sample restrictions, and measurement choices. Moreover, the results are unique to suicide deaths and not present for other causes of mortality.

Based on estimates that place the modal suicide victim in the middle of income distribution (consistent with the modal members of the general population), we interpret the opposite and significant impacts of upper and lower tail inequality as supportive of the notion that individuals look to others when evaluating their own utility and, in fact, look both ways, caring about the incomes of those above and those below them. While our semi-aggregate data do not allow us to directly verify this view, we are able

to confirm it using a individual-level data set constructed from actual information on geographic location, demographics, and other controls from the death records combined with estimates of individual income obtained from a nonparametric matching process applied to decennial Census data.

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 conclusions we draw are in Section 6.

2. Empirical Tests of the Relative Income Hypothesis

2.1 Existing Literature

Although the idea that relative income can affect utility is more than a century old, the empirical literature on the topic has emerged only recently. 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 subjects' negative reactions to the presence of a hierarchy, i.e., "inequality aversion," and/or 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 and contrived nature of these experiments make their results difficult to generalize.

The second vein of empirical work on relative income exploits information gathered from surveys of happiness and/or life satisfaction. Research using these data focuses on the relationship between happiness and relative position, generally measured as relative consumption, wages, or income. Along

these lines researchers have looked at relative wages and job satisfaction (Clark and Oswald 1996;

Brown, Gardner, Oswald, and Qian 2004), and individual happiness and local, regional, and national

income (Helliwell 2002; Luttmer 2005; Tomes 1986; and Ferrer-i-Carbonell 2005). Like those from the

experimental literature, results from investigations of happiness data lend support to the relative income

hypothesis. However, concerns about the quality of data on self-reported happiness (Brekke 1997 and

Osmani 1993; see Bertrand and Mullainathan 2001 for a broader critique of subjective survey data) have

limited the broad acceptance of results on relative income and interpersonal comparisons. These

concerns, involving language ambiguities (respondents may not all agree on the exact meaning of terms

like "happiness" and "life satisfaction"), scale comparability (is one person's "very satisfied" higher,

lower, or equal to another person's "satisfied"?), ambiguity regarding the time period over which

respondents base their answers, and respondent candidness, are ones overcome in data on suicide deaths.

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 revealed choices is thought to be preferable to relying on individual self-reports of

their feelings or 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,

---

[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]Here we are assuming that suicide decisions are not always the result of irrational and impulsive actions.
If our assumption fails to hold, i.e., suicide is a totally irrational and impulsive action, we should find no affect of
relative income (or any economic factor) in our analysis.

covering the universe of reported suicides by year.[5]

Despite these advantages, suicide data are not perfect. Suicide victims presumably are at the extreme tail of the distribution of life satisfaction over the population, and 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 we should find no correlation between relative income and suicide risk–a non-finding.

It is important to note that we are not the first to consider how economic variables may influence suicide risk. Hamermesh and Soss (1974) 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 (2000) considers suicide as one of several causes of death and finds that, unlike other negative health outcomes which 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. In a related survey article on subjective well-being and suicide risk, Oswald (1997) finds that many variables positively (negatively) associated with reported happiness are negatively (positively) associated with suicide risk.[6] To our knowledge we are the first to use information on suicide

---

[5]Note, however, that reported suicides are likely an undercount of 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 1999). In Section 5.3 we confirm that our empirical results are insensitive to this potential misclassification.

[6]Other recent examples of economists trying to explain suicide behavior include Cutler, Glaeser, and Norberg (2000), who attempt to explain the rise in teenage suicide and Brainerd (2001), who documents a strong link between macroeconomic instability and male suicide in the transition economies of the former Soviet Union during the 1990s. See also Marcotte (2003), Stevenson and Wolfers (2000), Chuang and Huang (1997), Huang (1996), Kimenyi and Shughart (1986), Hamermesh (1974), and Schapiro and Ahlburg (1982-83).

risk to study the existence and nature of interpersonal comparison.

3. Framework for Analysis

Although our empirical analysis is done at the group level, our underlying intuition is based on

individual behavior. In this section we describe our intuition and how it fits well into the existing

framework of random utility models. We then lay out the empirical formulation and estimation strategy.

3.1 Intuition and Basic Model

We begin with the premise that all individuals regularly assess the value of future life against the

current value of exit. For the vast majority of the population, this evaluation never nears the threshold

defining exit but rather determines the extent to which someone experiences being happy or not so happy.

For the marginal person, the threshold is binding and the evaluation ends in suicide. Assuming that the

factors affecting this continuum of happiness can be evaluated using the subset of individuals who

actually cross the suicide threshold (the marginal individual) we can use data on suicide deaths to directly

test hypotheses related to the interpersonal comparisons.

This intuition fits nicely into the framework of a random utility model (RUM).[7] See Appendix A

for a complete model of suicide choice based on the RUM. Here, we lay out the intuition of the model in

order to illustrate how it leads to empirical specifications at both the individual and semi-aggregate level.

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 a unobserved idiosyncratic component θ_i, which can

be treated as a random variable.[8] The deterministic component $u()$ depends on various state variables

---

[7]Our formulation of the random utility model happens to be quite consistent with research in positive psychology Peterson, Park, and Seligman (2005) and behavioral genetics (Lykken, 1999) which argues that happiness is a function of an individual-specific set point, exogenous life conditions, and choice variables like volunteering and leisure time.

[8] θ_i is analogous to the set point in the positive psychology models.

contained in the vector **x**. A key parameter in this model is θ_i, an unobserved, individual-specific component that incorporates all possible exogenous risk factors that determine an individual's predisposition to commit suicide. Higher 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.[9] More generally, the individual-specific component of the RUM 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** – this will be important when we specify our semiaggregated econometric model.

Within the structure of a RUM, we think of individuals choosing whether or not to commit suicide as a process by which they compare the utility they will receive from exiting today to the present value of expected future utility. This decision can be described by a very general model under two basic assumptions: (1) individuals weigh expected future utility against current exit utility when making the decision as to whether to commit suicide, and (2) there is a weakly exogenous (i.e., predetermined) "predisposition" factor, θ_i, that gives one's propensity to commit suicide for a given **x**. The general model requires no assumptions regarding what determines this predisposition factor.

Under these assumptions, the probability that an individual commits suicide is simply the

---

[9]For example, neurologists have found that suicidal patients and suicide victims tend to have particularly low cerebrospinal fluid levels of 5-hydroxyindoleacetic acid (5-HIAA), a metabolite of serotonin (serotonin transmits nerve impulses across synapses). Low levels of 5-HIAA also have been found in perpetrators of murder, arson, and other violent behavior linked to impulsiveness (see Hendin 1995). Such characteristics of the individual's brain are thus captured by θ_i. Also in θ_i would be psychological factors such as manic depression.

probability that his/her θ_i is above some threshold level $\theta^*(\mathbf{x}_i)$. If $\theta_i \sim G[\ ]$, then

$\Pr[\theta_i > \theta^*(\mathbf{x}_i)] = 1 - G[\theta^*(\mathbf{x}_i)] = f(\mathbf{x}_i)$. Parameterizing $f(\ )$ (for example, by assuming the suicide

probability is linear in $\mathbf{x}$), a researcher equipped with individual level data on suicide and the variables in

$\mathbf{x}$ could estimate the average effect of any variable in x on the likelihood of suicide, and thus infer the

effect of the variable on utility.[10]

3.2 Empirical Specification–Group-level Analysis

Since we do not have a complete data set at the individual level we must develop an econometric

model at the group-level. We observe cells (age×race×sex×county×year) with information on suicide

rates and a host of control $\mathbf{x}$ variables. To move from the individual model described above to the group-

level specification we estimate we need to assume that the variables in $\mathbf{x}$ are constant across members

within a group.[11] Taken this as given means that if $\mathbf{x}_i = \mathbf{x}_c$, the group level aggregate value of x, then the

probability of suicide within a group is simply $\Pr[\theta_i > \theta^*(\mathbf{x}_c)] = 1 - G[\theta^*(\mathbf{x}_c)] = f(\mathbf{x}_c)$, and the suicide rate

for a group is:

$$r_c = f(\mathbf{x}_c) + u_c. \qquad\qquad (1)$$

Linearizing (1), we obtain the basis for the primary estimations we perform in this paper. Equation (1)

can be estimated via OLS based on the assumption that the predisposition factors, now reflected in u_c, are

orthogonal to the variables in $\mathbf{x}_c$.

Of course, the assumption $\mathbf{x}_i = \mathbf{x}_c$ is unlikely to hold exactly. Realistically, $\mathbf{x}_i$ will vary to some

extent within each group c. As the within-group variance of $\mathbf{x}_i$ increases, our ability to infer individual

behavior from the group behavior– the ecological inference problem – will worsen. For example, holding

a group's mean income constant, the higher the group's variance of income, the larger the share of

---

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

[11]In this case the parameters of the individual model (e.g., the parameters of $f(\ )$) can be estimated using the group level data (see. e.g., King & Rosen 1995).

-8-

members that will have income below any particular level. If this unobserved share of low income members contributes to the suicide rate and is correlated with observed variables in $\mathbf{x}_c$, then the ordinary least squares estimator will be inconsistent.[12] We address this general concern in Section 5.3 by constructing an, albeit imperfect, individual level data set to evaluate whether the results we obtain using group-level data are driven by these type of ecological inference problems. We find the individual level estimations support the group level results.[13]

4. Data Sets and Variable Specification

The data used in this paper primarily come from two sources, the National Center for Health Statistics (NCHS)–for the mortality data, and the U.S. Census Bureau–for local area data such as income and income inequality. The NCHS provides two different data sets, the comprehensive but semi-aggregate Compressed Mortality Files (CMF) and the less comprehensive but individual-level Detailed Mortality Files (DMF). The empirical analysis in this paper mainly focuses on the former, though we provide some supporting results based on the DMF data. The following section briefly describes the construction of our primary data set, based on the CMF data. A more detailed description is supplied in Appendix B. Section 4.2 describes the individual level data set we constructed from the DMF.

4.1 Main, Semi-Aggregate Data

The data used in our main empirical analyses are two semi-aggregate data sets (1990 and 2000)

---

[12]We are able to handle the specific concern of this example in some of our regressions by controlling for the poverty rate. Note, however, that this is not a perfect control because (1) the poverty rate does not vary across cells within a county and (2) the poverty *level* is defined nationally, not locally.

[13]Ideally we would want to estimate panel data treating θ_c as a fixed effect and allowing us to control for possible correlation between θ_c and $\mathbf{x}_c$. The key hurdle for estimating a panel based on our data is that a number of key (time-varying) control variables are not available at the disaggregate level on an annual basis.

on suicide rates, known correlates of suicide, and measures of absolute and relative income.[14,15] The data construction requires combining information from several different sources. Below we describe the data construction and variable definitions. The CMF data contain disaggregate counts and rates of all deaths recorded in the United States (including deaths for which individual information is not publicly available) from 1989 through 2002. We use these data to construct suicide rates by gender-age-race-county cells in the U.S. in 1990 and 2000, measuring age in 10-year reference groups and race as white, black, and other.[16] We also compute the nonsuicide mortality rate by cell to use as a control variable in our analysis. Excluding the nonworking age population we end up with an analysis data file containing about 90,000 cells (approximately 3,000 counties $\times$ 3 races (white, black, other) $\times$ 2 genders $\times$ 5 ages = 90,000 cells).

Our control and income variables come primarily from a second source, the Integrated Public Use Microdata Series (IPUMS). IPUMS provides individual level records based on a 5 percent sample of each decennial census. Because the geographic identifier in IPUMS is not the county, as in the CMF, but a Census-constructed area known as a PUMA (Public Use Micro Area) we constructed a bridge geographic variable (separately for 1990 and 2000), which we call "geo," identifying the smallest

---

[14]Ideally we would like a data set with income, location, age, race, gender, marital status, veteran status, etc. (everything potentially in **x**) over some time period at the individual level for all suicide victims as well as for the universe, or at least a representative sample, of nonsuicide victims. With such a data set, it would be straightforward to estimate the effect on the suicide decision of each potential factor using probit or logit regression analysis.

[15]The constraints on our data are as follows. The individual-level data from death records which contain information on location and certain demographic characteristics do not include information on income. In addition, the available individual level data on the overall population (e.g., from the Current Population Survey or the IPUMS samples of Decennial Census responses) do not provide mortality information. We are aware of only two data sources that contain both types of information–the National Health Interview Survey (far too few observations of suicide deaths) and the public-use National Longitudinal Mortality Study (no geographic information).

[16]To mitigate the problem that these cell-specific suicide rates are frequently zero (suicide is, after all, a rare event), we include suicides committed in either 1989 or 1990 in our "1990" cross-section and include suicides occuring in either 1999 or 2000 in our "2000" cross-section. Including only 1990 and only 2000 deaths in our analysis yielded very similar results to those presented here. Age is defined as a five year age group for those 20-24.

common area between pumas and counties.[17] In 1990, the geo matches to a single county (but many

puma's) in 22 percent of cases, to a single puma (but many counties) in 53 percent of cases, to a single

county and single puma in 23 percent of cases, and an aggregate of multiple counties and multiple pumas

in 2 percent of cases. This distribution is similar in 2000. For the remainder of the paper, we use the

terms "county" and "geo" interchangeably.

Using IPUMS, we construct three different types of variables. First, we compute cell-share

variables that give the proportion of a gender*age*race*geo cell with a particular attribute; share

variables are computed for marital status, hispanic, veteran status, vietnam veteran status, home

ownership, educational attainment, and white-blue collar occupation. We also compute the

unemployment rate and the average family size at the same cell level. These variables make up our vector

of controls in the empirical estimation.

Finally, we use the IPUMS to compute our main analysis variables–absolute and relative income.

In keeping with the bulk of the literature on interpersonal comparisons (e.g., Luttmer 2005 and Ferrer-i-

Carbonell 2004), we compute income and relative income at the county(geo)-level, implicitly assuming

that individuals' reference groups are locally geographic.[18] While this represents just one of several

reference groups one could test, it is the type of model most commonly tested in the literature. The ratios

we compute are the following: 90-10 (90th-percentile of income in the county divided by the 10th-

percentile of income in the county), 90-50, and 50-10. We use household rather than personal income as

our baseline under the assumption that individuals self-perceptions are formed based on their household

---

[17]PUMAs are contiguous areas constructed so that they have a population above, but as close as possible to 100,000. Since the geographic distribution of population changes over time, the 1990 puma's are not the same areas as the 2000 puma's. As such, the 1990 semiaggregated IPUMS data cannot be longitudinally linked to the 2000 data to form a panel data set. Thus, we use the data as repeated cross-sections. Ruggles, et al. 2004 provide details on the IPUMS data.

[18] This level of aggregation translates into a reference group that is countygeo-based rather than cell-based; in other words, by selecting this countygeo-aggregation for income we are asssuming that individuals look across all residents of their county rather than just at those with their same characteristics. Numerous reference groups are possible. We pursue the empirical significance of competing reference groups in other work.

socioeconomic status rather on their monetary contribution to that socioeconomic status. That said, for most of the analysis in the paper we check the robustness of the results substituting personal for household income; the patterns generally are the same.

While these variables compose our main analysis data, for the 1990 cross section, we are able to obtain additional county-level control variables from the Census Bureau's *USA Counties* data files. Using these files we obtain county-level data on a number of variables: growth in per capita income, variance of per capita income (over time), poverty rate, violent crime rate, population density, average commute time, share of population in professional occupations, share in blue collar occupations, share in psychiatric facilities, and the share in prison.[19] Finally, borrowing from the public policy literature on the effects of firearm availability, we use state-level data on the share of suicides committed via firearm. The data comes from Miller, Azrael, and Hemenway (2002) and refer to the period 1988-1997.

Our final data set contains 24,665 observations for 1990. This number comes from 908 geos $\times$ 3 races (white, black, other) $\times$ 2 genders $\times$ 5 ages $= 27,240$ cells, and then subtracting the 1,634 cells with no population and the 941 cells with undefined unemployment rates (i.e., with zero labor force). For 2000, we have 27,720 observations derived from 962 geos.

4.2 Auxiliary, Individual-Level Data with Estimated Income

In addition to the CMF-based data described above, we make use also of the Detailed Mortality Files (DMF), available from the Inter-university Consortium for Political and Social Research (ICPSR) (see U.S. Department of Health and Human Services (1992)). The DMF are essentially the data from all death certificates in the U.S.[20] However, for confidentially reasons, the county of the deceased in not reported in the public-use files for individuals from counties with population less than 100,000, which

---

[19]As of the time of this writing, these county-level variables are available for 1990 as well as past years; data beyond 1990 is not yet readily available.

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

account for about a quarter of U.S. counties. This feature of the DMF leads us to prefer the more

comprehensive CMF data described above for our main empirical analyses trying to identify the effects of

local area income and inequality on suicide risk. Nonetheless, the DMF data can be exploited both to

perform robustness checks on our main results and to provide an estimate of the income distribution of

suicide victims, which is vital to interpreting our main results.

For both of these purposes it is necessary to know the income of the suicide victims in the DMF.

Although income is not reported on death certificates and hence is not known for the suicide victims in

the data, we can estimate their incomes by matching them to persons in the PUMS data. To do this, we

combine the DMF data with the PUMS 5% sample of the 1990 census, which we described earlier. The

matching procedure works as follows: (1) for each suicide record, find all matching observations in

IPUMS, matching on county (geo), age, race, sex, hispanic status, education, and marital status; (2)

calculate average income (household and personal separately) for this matching cell; and (3) assign this

average income to the suicide observation. This procedure provides a reasonably accurate prediction of

income. An analysis of variance (not shown) shows that the matching variables jointly explain 24% of

the variation in actual income (across the roughly 7 millions observations in the PUMS). County is found

to have the greatest explanatory power, followed by education and then marital status.

With this procedure, we are able to estimate income for 57 percent of the suicide records from

1989-92, totaling 50,328 suicides (we use more than just 1990 records in order to maximize the number

of records used in our analyses).[21] We use the same matching procedure to obtain estimated income for

the non-suicide population (the PUMS sample). For comparability, we use this estimated income in the

---

[21]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 DMF 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 suicides. It should also be
noted that some death records include occupation and industry of the deceased, but not enough records contained
this information for us to usefully include these variables in our matching procedure.

empirical exercises below rather actual income.

5. Descriptive Statistics and Main Results

5.1 Descriptive Statistics: Suicide Risk and Model Variables

The U.S. suicide rate has been relatively constant since 1950, averaging about 12 per 100,000

persons (see WHO 2005).[22] Looking beneath this average, there is considerable variation in suicide risk

by gender, age, and race (Figure 1). Men are significantly more likely to commit suicide than are women;

this is true across all age groups (panels A and B).[23] Suicide rates vary slightly for men across various

age groups, with the young and old more likely to take their own lives than those in the middle-age range.

In contrast, suicide rates for women generally increase with age, although the suicide rate of 55-64 year

old women slipped below that of 35-54 year olds in 2000. Whites are more likely to commit suicide than

other races; this is true among both men and women (panels C and D). Among men, blacks are more

likely to commit suicide than non-whites of other races. The reverse holds for women.

There also is notable variation in suicide rates by region in the U.S. as displayed in panels E and

F of Figure 1. Suicide rates are lowest in the mid-Atlantic region and highest in the Mountain region of

the U.S.[24] Separate tabulations, not shown, point to additional variation within regions at the state and

even the county level. The variation in suicide risk across demographic groups and geographic areas

provides us sufficient variation to perform our analysis using semi-aggregated data by gender-age-race-

---

[22]From 1950 to 2000, the U.S. suicide rate has fluctuated within the narrow range of 10.4 to 13.5 per 100,000.

[23]In contrast, mental health scholars have found that women are twice as likely to *attempt* suicide as men (see, e.g., Colt 1991).

[24]Hamermesh and Soss (1974) noted that suicide rates during their sample period, 1947-67, were especially high in the states of the Pacific region. They attributed this to the "restless life often associated with this area." By 1990, though, suicide rates in this region were not much different from the national rates. Suicide rates in the Mountain region, however, were especially high in both 1990 and 2000. A number of researchers in the public health field have attributed this to the relatively high availability of firearms in the states of this region (see, e.g., Miller, Azrael, and Hemenway 2002).

county cells.

The above patterns are based on the comprehensive CMF data set. In order to get some idea of the income distribution of suicide victims, we now turn to the less comprehensive but individual-level DMF-based data set described in Section 4.2 above. Figure 2 plots the distribution of predicted income for working-age suicide victims in our sample against that of the general U.S. working-age population.[25] As the figure shows, the predicted income distribution of suicide victims (controlling for demographics) sits slightly left of the distribution for the general population.[26] However, the shape of the distribution is quite similar and importantly for our interpretation of the regression results below, the bulk of the distribution of suicide victims is around the middle of the general income distribution.[27] We take this as supporting evidence for the notion that middle-income persons, who make up the bulk of the mass of both the suicide population (numerator of suicide rate) and general population (denominator of suicide rate), likely drive variation in the suicide rate, thus allowing our semi-aggregate results below to be interpreted as reflecting the behavior of middle-income persons.

Turning to the explanatory variables in the model, Table 1 reports descriptive statistics for each variable and Table 2 reports the partial correlation of these variables with the suicide rate, controlling for gender, age, and race. We begin with the partial correlations to highlight the similarity between patterns found in our semiaggregated data and those found by others using individual-level information. These correlations then represent the first test of our data and their ability to shed any light on the individual-

---

[25]In order to make the comparison over the same geographic area, the general population distribution excludes individuals from the sparsely-populated counties which are not included in our DMF sample.

[26]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.

[27]Ideally, one could estimate the income distributions for each and every county (geo) in order to assess whether suicide victims in a given county tend to be massed around that county's median income. However, there are far too few suicide records to be able to reasonably estimate income distributions at such a disaggregate level. We argue that the aggregate relationship between in the suicide income distribution and the general distribution is generally likely to hold at the county level as well.

level questions we have posed. The partial correlations confirm several patterns found in the literature on

individual suicide risk. For example, using our semi-aggregated data we find a positive and significant

correlation between cell-specific suicide rates and veterans share, divorced share, and the unemployment

rate (see Kposawa 2001; Charlton, et al. 1992; and Blakely, et al. 2003). Similarly, suicide rates in a cell

are negatively correlated with the share of college educated persons and the share of persons that are

married. Although rudimentary, these correlations suggest that the semi-aggregated data constructed for

our analysis produce patterns consistent with the "stylized facts" regarding suicide reported in the

epidemiology/public health, psychology, and sociology literatures.[28]

Looking at the income variables, the partial correlations are suggestive of a relationship between

cell-specific suicide rates, median income, and each of our three measures of income inequality—the

90/10, 90/50, and 50/10 interpercentile income ratios. The correlations show a negative relationship

between median income and suicide risk. Interestingly, each of the income inequality measures have a

significantly positive relationship with suicide rates in 1990, but a significantly negative relationship in

2000. It is these relationships that we will focus on in the remainder of our analysis, and we will find that

the discrepancy between 1990 and 2000 does not survive multivariate analysis.

5.2 Baseline Regression Results

Our baseline regression, motivated by equation (1), is a cross-sectional weighted GLS regression

of the suicide rate by gender×age×race×county cells on a full complement of control variables, county

median income, and measures of relative income measured as the ratio of the 90th to the 10th, 90th to the

50th, and 50th to the 10th percentiles of the distribution of household income. Strong correlations among

our relative income measures necessitates separate estimation for the 90/10 and for the joint entry of the

---

[28]Partial correlations of the type reported here are actually quite prevalent in these literatures, often even the standard way of summarizing the data. Our multivariate results show that some of these "stylized facts" are not robust to the inclusion of additional variables.

90/50 and 50/10.[29] Observations are weighted by cell population.

The results are shown in Table 3. (Separate results for males and females, which are qualitatively similar to the full sample results, are shown in Tables A2 and A3). Standard errors, shown in parentheses, are obtained from a Huber-White heteroskedasticity-consistent VC matrix allowing correlation of errors within state clusters. Following the patterns established in the first two tables, the rows, showing coefficients and standard errors for each independent variable, are broken into three sections (for presentation purposes) based on the level of disaggregation of the variable.

Before discussing the results on relative income, it is worth reviewing the patterns for some of the control variables. The coefficients on the race and age variables mostly follow the patterns observed in the descriptive statistics. Even after controlling for various socioeconomic measures and other observables, suicide risk is highest for whites, followed by other, non-black races, and then blacks. The coefficient estimates suggest that, after controlling for other factors such as education and income, the suicide rate for blacks is between 10 and 15 per 100,000 lower than for whites and other races. Over these samples of working age population (20-64), suicide rises with age for women and is more or less u-shaped across the age distribution for men. Suicide risk across gender×age×race×county cells rises with the share of divorced individuals in that cell (the omitted category in the regression). All other marital statuses reduce suicide risk, with the largest reduction corresponding to the married and widowed statuses. Suicide risk consistently is highest among cells with higher concentrations of high-school educated individuals; risk falls as the share of college educated individuals in the cell increases. Availability of firearms at the state-level, as measured by the share of suicides committed with a firearm (the standard method of measuring firearm availability in the public health literature, see Miller, et al. 2002), is strongly positively associated with the suicide rate in both cross-sections. These patterns in the control variables are common across all of the regressions we discuss.

---

[29]Correlations among the income and income inequality measures are shown in Table A1.

Regarding the effect of the unemployment rate on suicide, we find a large, statistically significant and positive (deleterious) effect for males in both 1990 and 2000; the effect for females is smaller and is statistically significant only in 2000.[30] The coefficient estimates suggest that an increase of one percentage point in the unemployment rate increases the male suicide rate by about 26 per 100,000, which is more than double the aggregate male suicide rate. These results are consistent with the simple partial correlations in Table 2 and with prior studies focusing narrowly on the unemployment-suicide relationship.[31]

It is interesting to note that the coefficients we obtain on these control variables are remarkably consistent with findings based on subjective well-being data. In particular, Oswald (1997), in summarizing the findings of this literature, states that reported happiness is higher for married persons and lower for divorced persons, is increasing with education, and is much lower for the unemployed. On the other hand, our results regarding suicide diverge from those in the subjective happiness literature for age (U-shaped with reported happiness) and race (whites have higher reported happiness). These differences may imply variation in the shapes of the happiness and suicide distributions across age and race groups.

The variables that make up the focus of our analysis, median income and measures of relative income, are reported near the bottom of Table 3. Consistent with our earlier partial correlations, we find group-level suicide rates are negatively associated with county median income. The negative relationship between suicide and income is consistent with results from the happiness literature which finds reported

---

[30]The unemployment effect for males is larger than for females in both 1990 and 2000, but the difference is statistically significant only in 1990.

[31]Hamermesh and Soss (1974) find a positive relationship between the male unemployment rate and the male suicide rate using U.S. aggregate time series data (1947-67); Schapiro and Ahlburg (1982-83) also find such a positive relationship using U.S. aggregate data. Most recently, Ruhm (2000) finds that the overall suicide rate is countercyclical using state-level panel data. At the individual level, Kposawa (2001) (using U.S. data) and Blakely, Collings, and Atkinson (2003) (using New Zealand data) find that, for a given survey year, the unemployed are more than twice as likely as the employed to commit suicide within the following three years.

happiness to be increasing with income (Oswald 1997, Luttmer 2005). This effect holds for both specifications of the model (with the 90/10 and with the 90/50 and 50/10) and for both time periods, 1990 and 2000. The results on overall inequality, captured by the 90/10 interpercentile ratio, are insignificant in 1990, suggesting little relationship between overall household income inequality in a county and suicide rates of gender×age×race cells within that county. The coefficient on the 90/10 is negative and significant in 2000, suggesting that for that year, higher overall inequality in a county was associated with higher suicide risk for cell groups in that county.

Columns 3 and 5 of the table present our key results. These columns show a breakout of overall inequality into upper tail inequality–defined as the 90th/50th interpercentile ratio–and lower tail inequality–defined as the 50th/10th interpercentile ratio. The results suggest that controlling for differences in median county income, both upper and lower tail inequality have independent and statistically significant effects on cell-specific suicide risk. Importantly, the results show that inequality in the upper tail (90/50 income ratio) raises suicide rates and inequality in the lower tail (50/10 income ratio) lowers suicide rates. This finding holds in both years of our sample. Specifically, coefficient estimates imply that a doubling of the 90/50 ratio is associated with an increase in the suicide rate of 2.5 per 100,000 in 1990 and 1.7 per 100,000 in 2000. Compared with those magnitudes, the negative effect on suicide risk of an increase in the 50/10 is about 3 times smaller in 1990 and roughly equal in 2000.[32]

Also worth noting is the relatively high value of the R^2's obtained, which on average imply that over 40 percent of the variation in suicide rates across county-age-race-sex groups is explained by these independent variables. These high values suggest that there is sufficient systematic variation in suicide to make it a useful variable for analysis. Additional analyses (not shown) show that the demographic

---

[32]Another way of interpreting the magnitudes of these effects is by asking how much the suicide rate changes in response to a one standard deviation change in the 90/50 or 50/10 (these standard deviations are shown in Table 1). A one standard deviation change in the 90/50 is associated with an increase in the suicide rate (per 100,000) of about 0.5 in 1990 and 0.4 in 2000. A one standard deviation change in the 50/10 is associated with an decrease in the suicide rate of about 0.5 in 1990 and 0.8 in 2000.

dummy variables alone explain about 37 percent of the variation in suicide rates (in 1990). Adding the

(non-income) control variables explains an additional 5 percent, and the income variables add another 1

percent.

The baseline results provide very limited evidence of a relationship between overall inequality

(90/10 measure) and cell-specific suicide rates, suggesting that individuals are not averse to income

dispersion *per se*.[33] That is, the results are inconsistent with a model of general inequality aversion.[34] The

findings for the 90/50 and 50/10 are consistent with the hypothesis that utility is a function of both

absolute and relative income. Our estimated income distribution for suicide victims, described in Section

4 above, implies that, on net, the baseline results are driven by behavior of individuals at or near the

middle of the income distribution. This suggests that the upward 90/50 effect and downward 50/10 effect

on the suicide rate is evidence that relative income matters and in a symmetric fashion; at least the median

individual looks up and down the income distribution when evaluating utility, and feels worse when

he/she falls behind (keeping up with the Joneses) and better when he/she pulls ahead (staying ahead of the

Smiths). It is important to note that from this model we cannot say *why* relative income matters to utility.

It may be that the income of others (and in turn the utility of others) enters the utility function directly –

i.e., interdependent preferences – or it could be that relative income enters utility indirectly as a proxy for

something individual such as own income risk. These possible interpretations are discussed in Section 6.

Before interpreting our results we test the statistical robustness of the relationships found above

between the suicide rate and upper- and lower-tail inequality and consider other explanations for our

findings. We also consider whether our results are artifacts of aggregation and thus not representative of

individual behavioral responses. Sections 5.3, 5.4, and 5.5 report on the findings from these various tests.

---

[33]In other unreported regressions, we also have found that replacing the 90/10 interpercentile ratio with the Gini coefficient, as an alternative measure of overall income inequality, does not change this conclusion.

[34]Our results have nothing to say about "unfairness aversion," which, lamentably, sometimes is referred to as inequality aversion in other studies.

5.3 Statistical Robustness Checks

To ensure that the baseline findings are statistically robust, we compare results across different

econometric specifications, data and sample restrictions, and the inclusion of additional explanatory

variables available for the 1990 sample. For ease of presentation we report results only for our key

variables: median income and the measures of income inequality.[35]

The first panel of Table 4 reports results from our baseline regression including state fixed

effects.[36] Qualitatively, the inclusion of state fixed effects does not alter our results, though the

magnitudes of the coefficient estimates do decline Median income remains a negative and statistically

significant influence on suicide risk. The coefficients on the 90/50 and 50/10 continue to be of opposite

sign in both periods, and are statistically significant in three of the four cases.

The second panel of Table 4 show results from an negative binomial poisson estimation of

suicide counts by gender*age*rage*county cells, with cell population as the exposure and controlling for

the same set of independent variables as in Table 3. The poisson MLE estimator is a common method for

analyzing count data, and the negative binomial poisson (NBP) is typically used when there is a

preponderance of zeros. Again, the results are not qualitatively different than those from our baseline

---

[35]The patterns for our control variables change little in these alternative specifications of the estimation model. Full results are available from the authors upon request.

[36]A priori, it is not clear whether including state fixed effects is appropriate. On the one hand, the inclusion of these fixed effects puts considerable pressure on the data, especially given the preponderance of zeros for the cell-specific suicide rates, and we have attempted to control for measurable conditions across states and counties that may be correlated with our dependent and independent variables. On the other hand, we have no controls for access to mental health or other health care treatments and since access to these services may vary by state including state fixed effects represents an important check of our results. For example, Daniel Hamermesh pointed out that the use of antidepressant drugs are likely taken differentially across states. This differential consumption of anti-depressants may be positively correlated with income and negatively correlated with suicide risk (though the 1990 results should be relatively unaffected by this concern as antidepressent usage was substantially lower at that time). A similar argument could be made at the county level, although policies affecting access to health care seem more likely to vary by state than by county. Moreover, our key income variables vary by county so this strategy would not work in those regressions.

specification. In the NBP estimation, coefficients greater than one imply a positive influence on suicide likelihood while those less than one indicate a negative influence. The coefficient on median income is less than one and significant, consistent with our baseline model.[37] Similarly the coefficients on upper and lower tail inequality (90/50 and 50/10) indicate opposing effects, with 90/50 inequality associated with higher suicide risk and 50/10 inequality associated with lower suicide risk. Both upper and lower tail measures of inequality are statistically significant in both 1990 and 2000.

The next two robustness checks concern characteristics of our data and sample. Panel three of the table reports estimates from a sample of cells that excludes the very upper tail of the distribution of suicide rates.[38] Excluding the top one percent of suicide rates from our sample does not materially affect the results. The signs, magnitudes, and significance patterns of the coefficients are basically unchanged, suggesting that the baseline findings are not the result of outlier cells. Our second check on the data and sample is presented in Panel 4 of the table and focuses on the potential influence of population weighting on our results. The check involves deleting small-population cells and running the baseline specification without population weighting. This adjustment has little effect on the results for 1990. In 2000 this adjustment eliminates the result on upper tail inequality but has little effect on either median income or 50/10 inequality.

The next panel takes account of the fact that in the United States, deaths are recorded as suicides if and only if compelling evidence of intent is established by the police or a coroner (e.g., discovery of a suicide note and/or evidence of intentional, self-inflicted wounds). As such, many epidemiologists, psychologists, and sociologists argue that suicide is undercounted in U.S. death records (see, e.g., Moyer, Boyle, and Pollack 1989; Rockett and Smith 1999; and Mohler and Earls 2001). Our understanding of

---

[37]Statistical significance of the NBP estimates is evaluated relative to one.

[38]We exclude observations above the 99th percentile (100 suicides per 100,000 in 1990 and 114 per 100,000 in 2000). We also excluded observations above the 95th percentile of the cell specific distribution of suicide rates. The results are similar.

the procedures regarding cause-of-death classification is that suicides lacking clear evidence as such

likely are classified as accidents which include deaths from causes such as motor vehicle crashes, drug

overdoses, and falls. Assuming that accidents do include a meaningful number of misclassified suicides

we estimate our baseline model with deaths from suicides and accidents combined. If accidents include

these underreported suicides, or more generally reflect the outcome of riskier behavior consistent with

one's view of the future value of life, we should see patterns on the inequality measures similar to those

in Table 3. The results for the combined accident-suicide death rate are quite consistent with those for

suicide alone. In fact, the magnitude and statistical significance of the 90/50 and 50/10 inequality

coefficients are even greater than those for the suicide-only regressions in Table 3. (The higher precision

likely owes to the greater unconditional cross-sectional variance of the accident-suicide death rate due to

the lower frequency of zeros).

Finally, we examine whether some omitted, albeit measurable, variable is responsible for the

results. For the 1990 sample we are able to incorporate data from the *U.S.A Counties* database (U.S.

Census Bureau 2005) on population density, income growth, and intertemporal income variance by

geocounty.[39] The last three panels of Table 4 report the results of including these variables (one by one)

in our baseline model. Including population density in the model does little to change the baseline results,

suggesting that the income and income inequality results are not due to urban/rural differences in suicide

rates and income and/or inequality. The final two panels look at one-year income growth and 10-year

(backward) income variance. Again, including these variables has little effect on our baseline findings.[40]

---

[39]We also estimated regressions including other county-level control variables for 1990. These variables
are the crime rate, the share of the population with a blue collar occupation, the share with a professional occupation,
share in a psychiatric institution, share in prison, and the poverty rate. The crime rate has a positive and statistically
significant effect on the suicide rate for females, but not for males; the other three variables were insignificant in all
regressions. Including these variables had no notable effect on the coefficient estimates or statistical significance of
the other variables, thus we excluded them in the presented results in order to facilate comparisons of the 1990
results to the 2000 results.

[40]We also ran the model with 10-year income growth without effect.

Lower and upper tail inequality consistently are of opposite signs and significant in each case.

Taken together the checks in Table 4 confirm our baseline findings that 90/50 and 50/10 measures of income inequality differentially affect cell-specific suicide rates.

5.4 Alternative Explanations

Next we ask whether our results are specific to suicide and thus, as we will argue, reflecting effects of relative income on utility, or are more generally observed for other causes of death and thus more likely related to some other variable such as access to emergency medical care. Panels A and B of Table 5 shows results from our baseline regression estimated on two alternative dependent variables: (1) all- cause mortality and (2) deaths from heart failure. All-cause mortality rates are weakly and inconsistently linked to median income, and strongly positively associated with inequality, however measured. The results on all-cause mortality are consistent with those found by other researchers (e.g., Ruhm 2000). Importantly, these results depart completely from those for the suicide rate.[41]

The second panel of the table reports results using death rates from heart failure as the dependent variable. Our use of heart attack deaths is meant to test whether our results on suicide risk really owe to differential 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. Overall, the results show there is much less of an association between median income, income inequality, and death rates from heart failure than we found for death rates from suicide. Only in the 1990 sample are any of our income variables statistically significant. In terms of relative income, the 90/50 and 50/10 have opposite signs to each other but also opposite signs to the suicide results, suggesting that

---

[41]Ruhm (2000) also finds that suicide differs from other types of mortality in regressions controlling for economic variables.

90/50 inequality decreases deaths from heart failure while 50/10 increases deaths from heart failure. Only the 90/50 measure in the 1990 sample is statistically significant. The results on all-cause and heart failure mortality support the idea that the relationships we find for suicide risk are entirely unique to suicide and most likely related to relative income concerns in individuals' utility functions.

We have shown that the suicide rates across very disaggregate cells covary systematically with a larger number of economic and non-economic variables. Moreover, they, and they covary in a manner consistent with what one would expect based on findings using subjective well-being data. Based on an assortment of robustness checks, the results appear not to be driven by omitted variable bias or by measurement error in our dependent variable. Parallel examinations of all-cause and heart failure mortality show that our findings on upper and lower tail income dispersion are unique to the suicide sample. And since suicide represents a very small share of any cell's population, it is unlikely that our results are driven by reverse causality. Thus, we conclude that the relationships we estimate between the income inequality measures and cell specific suicide rates represent actual effects of upper-tail and lower-tail inequality on suicide risk.

5.5 Artifact of Aggregation?

Although the semiaggregate results hold up to a variety of robustness checks it could be the case that the findings do not reflect underlying behavioral responses of individuals but rather are a misleading artifact of aggregation. To examine this possibility we turn to our estimated individual-level data described previously. Our objective is to estimate the effects of own income and own income relative to reference-group income on the probability of suicide, using probit and logit analyses. These estimations are analogous to recent individual-level analyses performed by others using happiness survey data (e.g., Luttmer 2005 or Ferrer-i-Carbonell 2004). We use them as a check on our aggregate results.

Using estimated individual data on income, we examine two basic models regressing individual suicide probability on own income, own income relative to reference-group (e.g., county) income, and a

set of control variables analogous to those in our earlier, semi-aggregate regressions.   The first has a

similar structure as the earlier regressions though at much more disaggregate level and using the

individual-level Detailed Mortality Files as opposed to the more comprehensive but semi-aggregate

Compressed Mortality Files. We compute suicide rates at the least aggregate level possible given the

variables available in the DMF–the level defined by county, age, race, sex, hispanic status, education, and

marital status. We use the number of individuals (times 20) in the IPUMS five-percent sample in each

cell as a (admittedly noisy) measure of cell population used in the denominator of the suicide rates. We

interpret these suicide rates as estimates of the suicide probability for individuals within these cells.

Combining these suicide rates with predicted income for these individuals (which varies at the

same level of aggregation) and control variables analogous to those in our earlier regressions (e.g., cell

unemployment rate, share of cell that are Vietnam War veterans, and share that own their home), we

regress the suicide probability on own (cell) income, average county income, key demographic dummies,

and other control variables. Note that in our regressions using the DMF-PUMS data, we exclude

dummies for the key income predictors used in our income estimation procedure, namely, county,

education, and marital status. Otherwise, due to mulicollinearity between these variables and predicted

income, there would be little independent variation with which to identify the coefficient of own income.

Our strategy in these regressions amounts to treating estimated income as a summary statistic for

socio-economic status.

The results of this regression are reported in Table 6. Consistent with the results from the semi-

aggregate analysis, we find that suicide risk falls as own income rises, but rises with the income of one's

county. This result is consistent with those from our semi-aggregate regressions.

The OLS regression is nearly equivalent to combining the suicide records (with predicted

income) with the individual records from IPUMS and performing an individual-level probit or logit

estimation of the probability of commiting suicide as a function of estimated own income, average county

income, and controls. This is particularly true if one clusters on cell and accounts for the fact that each

IPUMS individual represents 20 individuals from the wider population.[42] Panel B of Table 6 reports

results from such a probit regression (logit estimation yielded nearly identical results). The coefficients

represent the marginal effects evaluated at the mean. Again, the coefficients are of the expected sign and

are significant.[43]

Importantly, these individual-level results are not only consistent with our earlier, semi-aggregate

results, but also are remarkably consistent with recent empirical analyses using subjective well-being

survey data. Most notably, Luttmer (2005) matches data on self-reported happiness (7-point scale),

household income, and demographic controls for about 10,000 individuals to average income data at the

PUMA level. He finds that, controlling for own income, PUMA income is negatively associated with

happiness.[44]

6. Assessment of the Results and Discussion

Using a unique data set on suicide risk, we find compelling evidence in support of the idea that

individuals care about their own income and the income of others around them. These results are

consistent with the findings of other studies relying on data from happiness surveys and/or controlled

experiments. These findings, from a baseline regression of cell-specific suicide rates and relative income,

are remarkably resilient to changes in model specification, sample selection, and to the inclusion of

additional variables related to income. We also find that these findings are unique to suicide as a cause of

---

[42]For the sake of robustness, we in fact cluster on more than cell. We cluster at the level of county.

[43]Note that for the probit results in Table 6 we display the p-values of the coefficient estimates (with respect to the null that the coefficients are equal to zero) instead of standard errors as the latter are extremely small and hence difficult to display.

[44]Ferrer-i-Carbonell (2005) obtains similar results using German data. Neither Ferrer-i-Carbonell nor Luttmer address the issue of asymmetric effects for those above and below reference-group income.

death, shedding doubt on explanations of the findings based on access to emergency or other medical care.

Based on estimates of income that show the bulk of the suicide victims come from the middle of the income distribution, we interpret our results as evidence that relative income, possibly in addition to absolute income, affects utility and does so in a two-sided manner. Utility rises as one's income pulls away from the bottom of the distribution in one's local area and falls as one moves away from the top of the local distribution. While this could reveal a form of interdependent preferences whereby individuals enviously look up at the Joneses and gratefully look down upon the Smiths, an alternative interpretation is that relative income proxies for income risk. If negative and positive shocks to income are equally likely but the distribution of shocks is proxied by the observed *ex-post* income distribution, then middle-income individuals may want a long distance between themselves and the bottom of the income distribution (e.g., the 10th percentile) – i.e., a low risk of falling to that bottom income level – but desire a very short distance between themselves and the most well off – i.e., a high probability of rising to that high income point. Both this risk story and the interdependent preferences hypothesis are consistent with the observed effect of relative income on utility. Distinguishing between these as well as other possible interpretations are subjects for future research.

Appendix A. An Individual Model of Suicide Choice

In Section 3.1, we described the basic intuition of the individual level model of suicide that forms

the theoretical framework guiding the empirical works we perform in this paper. This Appendix lays out

that model in detail.

Our model is based on the random utility model, which specifies that individual utility has the

form $u(\mathbf{x}_{it}) - \theta_i$. That is, contemporaneous utility for individual i in period t is a function of a

deterministic component, $u()$ and a unobserved random component θ_i. The deterministic component $u()$

depends on various state variables contained in the vector $\mathbf{x}$. As is standard, we assume $u()$ is monotonic

in each variable in $\mathbf{x}$, strictly concave, and twice-differentiable. Mathematically, this can be written as the

following dynamic programming problem:

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

subject to the boundary conditions

$$V(\mathbf{x}_{iT} \mid \theta_i) = 0; \ S_{it} = 1 \Rightarrow (S_{i,t+s} = 1) \ \forall \ s > 0 \ _{.45}$$

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

---

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

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

where $P(T - t)$ is the probability of surviving (by nonsuicide) to next period given the individual's maximum remaining lifespan, $T - t$. It should be the case that $P(0) = 0$, $P(T-t) \geq 0$, 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.

specific component of utility.[46] β is the discount factor and $E_t V(\mathbf{x}_{i,t+1})$ is the expected value of future

utility conditional on this period's decision, S_t. Hereafter, let V_{it} denote $V(\mathbf{x}_{it}|\theta_i)$.

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 $V_{i,t+1}^c$ 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:

$$S_{it}^* = \begin{cases} 1, \text{ if } u(x_{it}) - \theta_i + \beta E_t V_{it+1}^c < 0 \\ 0, \text{ otherwise} \end{cases}, \text{ where} \tag{2}$$

$$E_t V_{t+1}^c = E_t \left\{ \sum_{j=0}^{T-(t+1)} \beta^j \left[u_{t+1+j}(\mathbf{x}_{t+1+j}) - \theta_i \right] \left(1 - S_{t+1+j,t+j} \right) \right\}. \tag{3}$$

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 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.

---

[44]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 in the sympathy and attention one might receive posthumously, implying a positive value.

Conditional on $\mathbf{x}_{it}$ and β, whether $u(\mathbf{x}_{it}) - \theta_i + \beta E_t V_{it+1}^c(\mathbf{x}_{it} \mid \theta_i)$ is positive or negative will be

completely determined by the value of the idiosyncratic component θ_i. For a given $\mathbf{x}_{it}$, there exists a

critical value $\theta^*(\mathbf{x}_{it})$ above which $u(\mathbf{x}_{it}) - \theta_i + \beta E_t V_{it+1}^c(\mathbf{x}_{it} \mid \theta_i)$ 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 $\theta^*(\mathbf{x}_{it})$ with respect to that variable using the implicit function theorem[47]:

$$\frac{\partial \theta_t^*}{\partial x_{it}} = - \left[\underbrace{\frac{\partial u_t}{\partial x_{it}}}_{A} + \beta \underbrace{\frac{\partial E_t V_{t+1}^c}{\partial x_{it}}}_{B} \right] \Big/ \frac{\partial F}{\partial \theta_t^*}, \tag{4}$$

where from (3) we get

$$\frac{\partial E_t V_{it+1}^c}{\partial x_{it}} = \sum_{j=0}^{T-(t+1)} \beta^j \left[\frac{\partial u_{t+1+j}}{\partial x_{it}} \left(1 - S_{i,t+1+j,t+j}\right) - \frac{\partial S_{i,t+1+j,t+j}}{\partial x_{it}} u_{i,t+1+j} \right].$$

By construction, $\partial F / \partial \theta_t^* < 0$ (cancelling out the negative sign in front of the brackets in (4)), so the sign

of $\partial \theta_t^* / \partial x_{it}$ will equal the sign 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**; we call this the signal effect. For the variables considered in the empirical analysis of this

paper, the signal effect is likely to be near zero and hence the overall effect will be dominated by the

contemporaneous effect. Nonetheless, it is important to recognize that effects of variables on utility and

---

[45]The implicit function is the solution for θ^* of the equation:

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

suicide that we identify represent the *net* effect of these two countervailing forces.[48]

---

[46]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: Data Construction

Ingredient 1: Suicide and Nonsuicide Mortality Rates by County-Race-Sex-Age

The binding constraint in this data exercise is population data, which is necessary for the denominator of any by-cause mortality rate. The narrowest possible level of aggregation for U.S. population data from the Census Bureau is a county-race-sex-age cell.[49] Suicide frequencies at the same level can be computed from individual death records. The Compressed Mortality Files (CMF), constructed by the National Center for Health Statistics, contains these frequencies along with the Census population data. The CMF data cover all deaths occurring in the United States (including deaths for which individual information is not publicly available), currently for years between 1989 and 2002. From this data, we constructed suicide rates for all county-race-sex-age cells in the U.S. in 1990 and 2000.[50] Age here refers to a 10-year age group (except the 5-year group 20-24). Since mortality risk may be a factor in the suicide decision (as could be captured in our model by the $P(T\text{-}t)$ term in footnote 6), we also construct the nonsuicide mortality rate at the same cell level. Excluding the non-working-age population (< 20 or > 64), there are roughly 3,000 counties $\times$ 3 races (white, black, other) $\times$ 2 genders $\times$ 5 ages $=$ 90,000 cells.

---

[49]An alternative to using Census population data, which would allow for an even more disaggregate cell (e.g., by education as well) would be to impute population based on the IPUMS 5 percent sample of the 1990 decennial census. However, the population estimates would be frequently be based on extremely small sample cells and hence would be extremely noisy. Keeping these concerns in mind, we nonetheless do exactly such an imputation later in the paper as a check on whether our main results could be driven by aggregation bias/ecological fallacy.

[50]To mitigate the problem that these cell-specific suicide rates are frequently zero (suicide is, after all, a rare event), we include suicides occuring in either 1989 or 1990 in our "1990" cross-section and include suicides occuring in either 1999 or 2000 in our "2000" cross-section. Including only 1990 and only 2000 deaths in our analysis yielded very similar results to those presented here.

Ingredient 2: Income by County and Controls by County-Race-Sex-Age

We construct data at a similar level of disaggregation on income, demographics, and other control variables using the Integrated Public Use Microdata Series (IPUMS) (see Ruggles, et al. 2004). Among other data, IPUMS provides individual level records based on a 5 percent sample of each decennial census. However, the geographic identifier in IPUMS is not county but rather an Census-constructed area called a Public Use Micro Area (puma). Pumas are contiguous areas constructed so that they have a population above, but as close to, 100,000 as possible.[51] In some cases, there is a one-to-one mapping between county and puma; in some cases, multiple neighboring small counties are pooled together into a single puma; and in some cases, a single large county is broken up into multiple pumas. In a small number of cases, part of a county may go to one puma while other parts go to a different puma. For the purpose of merging the county-level suicide rate data with the puma-level IPUMS data, we constructed a new geographic variable (separately for 1990 and 2000), which we call "geo," identifying the smallest common area between pumas and counties. In 1990, the geo matches to a single county (but many puma's) in 22 percent of cases, to a single puma (but many counties) in 53 percent of cases, to a single county and puma in 23 percent of cases, and an aggregate of multiple counties and multiple pumas in 2 percent of cases. This distribution is similar in 2000. For the remainder of the paper, we will generally use the terms "county" and "geo" interchangeably.

From IPUMS, we are able to construct three different sets of variables. First, we constructed cell share variables, such as the proportion of the cell in each educational category at the county(geo)-race-sex-age cell level. Other such share variables captured the following characteristics: education, professional occupation, blue collar occupation, marital status, hispanic, veteran status, vietnam veteran status, and home ownership. Second, we computed the unemployment rate and the average family size at

---

[51]Since the geographic distribution of population changes over time, the 1990 puma's are not the same areas as the 2000 puma's. Therefore, in general, the 1990 semiaggregated IPUMS data cannot be longitudinally linked to the 2000 data. This leaves us with repeated cross sections, but not a panel data set.

the same cell level.

The third set of variables, county-level income interpercentile ratios that we construct from IPUMS, are the focus of this paper. The ratios we compute are the following: 90-10 (90th-percentile of income in the county divided by the 10th-percentile of income in the county), 90-50, and 50-10. We use household rather than personal income as our baseline under the assumption that individuals self-perceptions are formed based on their household socioeconomic status rather on their monetary contribution to that socioeconomic status. That said, for most of the analysis in the paper we check the robustness of the results substituting personal for household income; the patterns generally are the same. We use measures of income and income inequality at the *county* level for two reasons. First, we believe the most plausible definition of the reference group for which income comparisons are made is local area neighbors, regardless of their race, sex, or age; and county is the narrow geographic information available in the CMF. Second, practically speaking, income inequality measures calculated from a small number of individuals, as would be the case if we used groups below the county level (given the IPUMS data is already just a 5 percent sample), would be extremely noisy.

Ingredient 3: Other County- and State-Level Controls

For the 1990 cross section, we obtained additional county-level control variables from the Census Bureau's *USA Counties* data files. These files provided us county-level data on a number of variables: growth in per capita income, variance of per capita income (over time), poverty rate, violent crime rate, population density, average commute time, share of population in professional occupations, share in blue collar occupations, share in psychiatric facilities, and the share in prison.[52] Another potentially important control variable is accessibility of the typical "instruments" of suicide. The most common such instrument, especially for males, and also the one most likely to vary geographically, is firearms.

---

[52]As of the time of this writing, these county-level variables are available for 1990 as well as past years; data beyond 1990 is not yet readily available.

Borrowing from the public policy literature on the effects of firearm availability, we use state-level data on the share of suicides committed via firearm. We also try using the share of homicides committed via firearm; the results are similar as the two sets of shares are highly correlated. The data comes from Miller, Azrael, and Hemenway (2002) and refer to the period 1988-1997.

After merging these three data ingredients together, we arrive at 24,665 observations for 1990. This number comes from 908 geos $\times$ 3 races (white, black, other) $\times$ 2 genders $\times$ 5 ages = 27,240 cells, and then subtracting the 1,634 cells with no population and the 941 cells with undefined unemployment rates (i.e., with zero labor force). For 2000, we have 27,720 observations derived from 962 geos.

<h1 align="center">References</h1>

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

Abel, Andrew B. "The Effects of a Baby Boom on Stock Prices and Capital Accumulation in the Presence of Social Security." *Econometrica*, March 2003, 71(2), pp. 551-78.

Adler, Nancy E., W. Thomas Boyce, Margaret A. Chesney, Susan Folkman, and S. Leonard Syme. "Socioeconomic Inequalities in Health. No Easy Solution." *JAMA*, June 23-30, 1993, 269(24) pp. 3140-5.

Akerlof, George A. and Janet Yellen. "The Fair Wage-Effort Hypothesis and Unemployment," *Quarterly Journal of Economics*, May 1990, 105(2), pp. 255-83.

Alesina, Alberto, Rafael Di Tella, and Robert J. MacCulloch. "Inequality and Happiness: Are Europeans and Americans Different?" *Mimeo*, 2001.

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*, March 2005, 56(3), pp. 405-21.

Becker, Gary S. and Nigel Tomes. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, March 1986, 4(3), pp. S1-39.

Becker, Gary S. and Luis Rayo. "Evolutionary Efficiency and Happiness." Mimeo, University of Chicago, 2005.

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.

Boskin, Michael J. and Eytan Sheshinski. "Optimal Redistributive Taxation when Individual Welfare Depends upon Relative Income." *Quarterly Journal of Economics,* November 1978, 92(4), pp. 589-601.

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, no4-6, 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 BM de Waal. "Monkeys Reject Unequal Pay." *Nature*, September 18, 2003, 425, pp. 297-99.

Brown, Gordon D.A., Jonathan Gardner, Andrew Oswald, and Jing Qian. "Does Wage Rank Affect

Employees' Wellbeing?" *IZA Discussion Papers,* March 2005, 1505, Institute for the Study of Labor (IZA).

Chapman David A. "Does Intrinsic Habit Formation Actually Resolve the Equity Premium Puzzle?" *Review of Economic Dynamics*, 2002, 5(3), pp. 618-45.

Charlton, J., S. Kelly, K. Dunnell, B. Evans, R. Jenkins, and R. Wallis. "Trends in Suicide Deaths in England and Wales." *Population Trends*, 1992, 69 10-16.

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.

Colt, George Howe. *The Enigma of Suicide : A Timely Investigation into the Causes, the Possibilities for Prevention and the Paths to Healing*. New York: Scribner, 1991.

Constantinides, George M. "Habit Formation: A Resolution of the Equity Premium Puzzle." *Journal of Political Economy*, March 1990, 98(3), pp. 519-43.

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., Greg J Duncan, George A. Kaplan, and John W. Lynch. "Macro-to Micro Links in the Relation between Income Inequality and Mortality." *The Milbank Quarterly,* September 1998, 76(3), pp. 315-39.

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.

Di Tella, Rafael, Robert J. MacCulloch, and Andrew J. Oswald. "The Macroeconomics of Happiness." *Review of Economics and Statistics*, Forthcoming.

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

Duesenberry, James S. *Income, Saving, and the Theory of Consumer Behavior*. Cambridge, MA: Harvard University Press, 1952.

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.

Feldstein, Martin S. "Income Inequality and Poverty." *National Bureau of Economic Research Working Paper*, October 1998, W6770.

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.

Ferrer-i-Carbonell, Ada and Bernard M.S. van Praag, "Income Satisfaction Inequality and its Causes." September 2003, *IZA Discussion Paper* No. 854; Tinbergen Institute Working Paper No. TI 01-014/3.

Ferrer-i-Carbonell, A. and Frijters, P. "How Important is Methodology for the Estimates of the Determinants of Happiness?" *Economic Journal*, forthcoming.

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.

Haan, Mary N., George A. Kaplan, and S. Leonard Syme. "Socioeconomic Status and Health: Old Observations and New Thoughts." In: J. P. Bunker, D. S. Gomby, & B. H. Kehrer, (Eds.), *Pathways to Health: The Role of Social Factors*, 1989, pp.76-135. Menlo Park, CA., HJ Kaiser Family Foundation.

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." *National Bureau of Economic Research Working Paper*, 2002, 9065.

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.

Ireland, Norman J. "Status-seeking, income taxation and efficiency." *Journal of Public Economics*, 1998, 70, pp. 99-113.

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

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

King, Gary, Ori Rosen, and Martin A. Tanner. Ecological Inference. Cambridge: Cambridge University Press, 2004.

Kposowa, Augustine J. "Unemployment and Suicide: a Cohort Analysis of Social Factors Predicting Suicide in the US National Longitudinal Mortality Study." *Psychological Medicine*, 2001, 31, no. 1 127-38.

Layard, Richard. "Wages, Policy and the Redistribution of Income." The Colston Research Society Annual Lecture, in D. Collard, R. Lecomber and M. Slater (eds.), The Limits to Redistribution, Colston Society, 1980.

Ljungqvist, Lars and Harald Uhlig. "On Consumption Bunching under Campbell-Cochrane Habit Formation," *Working Paper Series in Economics and Finance*, 1999, 337, Stockholm School of Economics.

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

Lykken, David. *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.

Marcus, Eric. Why Suicide? San Francisco: Harper, 1996.

Miller, Douglas, and Christina Paxson. "Relative Income, Race, and Mortality." *Mimeo, Center for Health & Well-Being, Woodrow Wilson School of Public and International Affairs, Princeton University*, 2001.

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.

Moore, David E. and Mark D. Hayward. "Occupational Careers and Mortality of Elderly Men." *Demography*, February 1990, 27(1), pp. 31-53.

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.

Ng, Yew-Kwang. "Relative-Income Effects and the Appropriate Level of Public Expenditure." *Oxford Economic Papers*, June 1987, 39(2), pp. 293-300.

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. "Three Theorems on Inflation Taxes and Marginal Employment Subsidies" *Economic Journal,* 1984, 94(375), pp. 599-611.

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.

Steven Ruggles, 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.

Seidman, Laurence S. "A Phase-Down of Social Security: The Transition in a Life Cycle Growth Model." *National Tax Journal*, 1987, XXXIX, pp. 97-107.

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

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.

U.S. Census Bureau. USA Counties 1998, January 2005. Available online: http://censtats.census.gov/usa/usa.shtml.

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.

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

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

Figure 1. Suicide Rates by Age, Race, and Region
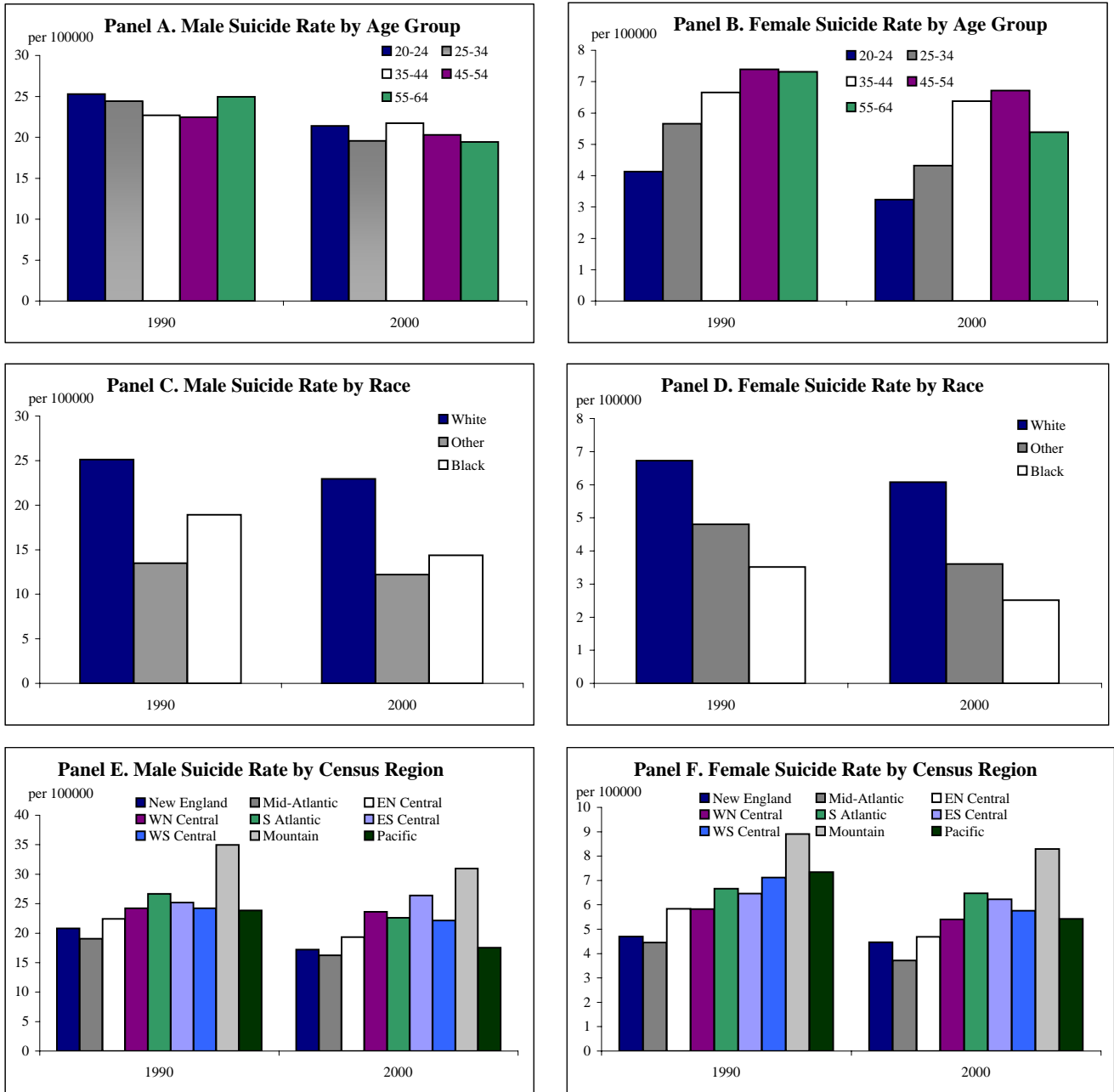
Source: Authors' calculations from NCHS Compressed Mortality Files.

Figure 2. Estimated Income Distribution for General and Suicide Populations, 1990
(Kernal Density Estimates)

	Suicide Population	General Population
Mean:	$36,863	$42,481
Median:	$33,768	$39,134
Std. Deviation:	$20,904	$20,281
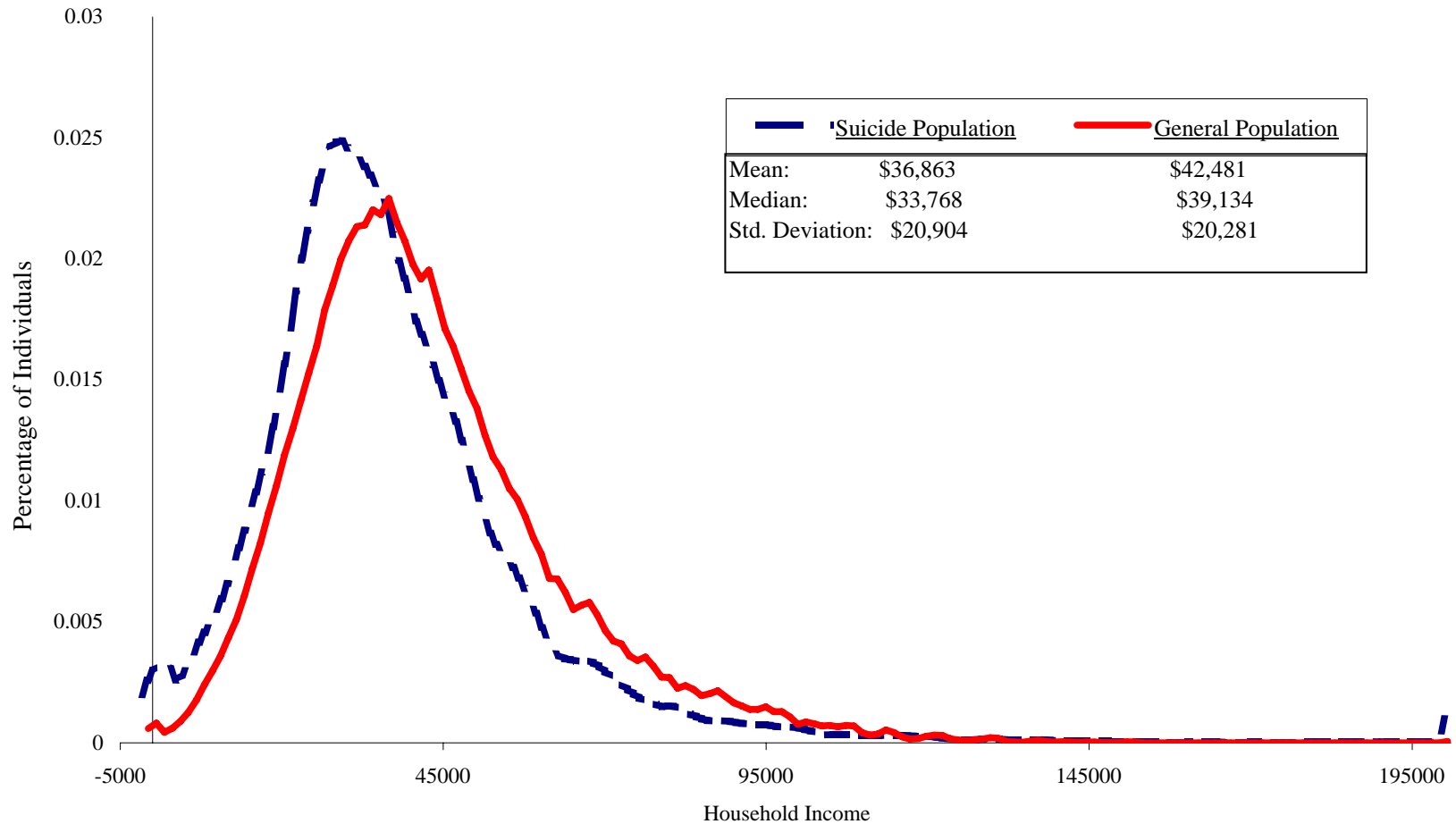
Notes: For both sample populations, individuals are assigned the mean income for their cell, where cell is defined by geo (county), age, race, sex, education, marital status, and hispanic status. Mean cell incomes are calculated using the IPUMS 5% sample of the 1990 Census. Individual observations for suicide population come from the Detailed Mortality Files for 1989-92; individual observations for general population come from the 1990 IPUMS 5% sample. Both populations exclude individuals from counties with population under 100,000, as such counties are unidentified in the Detailed Mortality Files.

Table 1. Characteristics of semi-aggregated data, by year

	1990		2000	
	Mean	S.D.	Mean	S.D.
Dependent Variable				
Suicide rate (per 100,000)	15.14	14.96	13.33	14.01
Independent Variables				
Dummy Variables				
Gender (Male=1)	0.49	0.50	0.50	0.50
White	0.85	0.36	0.82	0.38
Black	0.11	0.32	0.12	0.33
Other Race	0.04	0.19	0.05	0.23
Age 20 - 24	0.13	0.34	0.11	0.32
Age 25 - 34	0.30	0.46	0.24	0.43
Age 35 - 44	0.25	0.44	0.27	0.45
Age 45 - 54	0.17	0.38	0.23	0.42
Age 55 - 64	0.15	0.35	0.15	0.35
Vary by Gender, Age, Race, and Geo				
Non-suicide Mortality Rate (per 100,000)	353.67	438.17	315.84	358.21
Unemployment Rate	0.05	0.04	0.04	0.04
Average household size	3.03	0.49	2.85	0.49
Share Veteran	0.16	0.22	0.13	0.15
Share married	0.62	0.19	0.59	0.18
Share divorced	0.13	0.07	0.15	0.07
Share widowed	0.03	0.05	0.02	0.04
Share single /never married	0.22	0.23	0.24	0.23
Share with < 9th grade ed	0.06	0.06	0.04	0.05
Share with < 12th but > 8th grade ed	0.09	0.05	0.07	0.04
Share with 12th grade or GED ed	0.34	0.10	0.31	0.10
Share with some college	0.30	0.09	0.32	0.09
Share with college degree	0.15	0.08	0.17	0.08
Share with MA, prof degree, or PhD	0.07	0.06	0.09	0.07
Share Hispanic	0.07	0.13	0.09	0.15
Share of households that live in own home	0.68	0.18	0.72	0.21
Vary by Geo				
Median household income	38,005	8,741	45,045	10,691
90/10 ratio of household income	6.97	1.86	9.49	2.84
90/50 ratio of household income	2.14	0.21	2.48	0.32
50/10 ratio of household income	3.23	0.60	3.78	0.74
Vary by State				
Share of suicides committed by firearm	0.58	0.12	0.58	0.12
Number of Observations	24,665		27,720	

Notes: Unit of observation is a cell defined by gender*age*race*geo. Data are weighted by population.

Table 2: Suicide and key explanatory variables
Partial correlations controlling for gender, age, and race

	1990	2000
	Correlation	Corrrelation
Key Explanatory Variables		
<u>Vary by Gender, Age, Race, and Geo</u>		
Non-suicide Mortality Rate (per 100,000)	0.019 ***	0.053 ***
Unemployment Rate	0.074 ***	0.068 ***
Average household size	-0.044 ***	-0.022 ***
Share veteran	0.021 ***	0.068 ***
Share married	-0.050 ***	-0.014 **
Share divorced	0.128 ***	0.154 ***
Share widowed	0.009	0.030 ***
Share single /never married	-0.012 *	-0.074 ***
Share with < 9th grade ed	-0.012 *	-0.056 ***
Share with < 12th but > 8th grade ed	0.072 ***	0.109 ***
Share with 12th grade or GED ed	0.054 ***	0.139 ***
Share with some college	0.034 ***	0.049 ***
Share with college degree	-0.093 ***	-0.145 ***
Share with MA, prof degree, or PhD	-0.087 ***	-0.147 ***
Share Hispanic	-0.009	-0.104 ***
Share of households that live in own home	0.001	0.073 ***
<u>Vary by Geo</u>		
Median household income	-0.166 ***	-0.176 ***
90/10 ratio of household income	0.062 ***	-0.049 ***
90/50 ratio of household income	0.045 ***	-0.061 ***
50/10 ratio of household income	0.068 ***	-0.026 ***
<u>Vary by State</u>		
Share of suicides committed by firearm	0.168 ***	0.190 ***
Number of Observations	24,665	27,720

***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively

Table 3: Suicide risk and income inequality—results from baseline regression, by year

	1990				2000			
	Overall Inequality		Disaggregated Inequality		Overall Inequality		Disaggregated Inequality	
Control Variables								
Constant and Dummy Variables	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.
Constant	14.06 ***	(4.46)	10.40 **	(4.42)	13.95 ***	(5.26)	11.74 **	(4.98)
Male	16.26 ***	(0.97)	16.30 ***	(0.97)	15.57 ***	(1.03)	15.57 ***	(1.03)
Other Race	-3.03 **	(1.33)	-2.90 **	(1.34)	-1.32	(1.06)	-0.83	(1.05)
Black	-10.67 ***	(1.30)	-10.61 ***	(1.31)	-13.36 ***	(1.29)	-13.24	(1.28)
Age 25 - 34	5.28 ***	(0.86)	5.42 ***	(0.87)	5.38 ***	(0.91)	5.54 ***	(0.91)
Age 35 - 44	6.64 ***	(1.19)	6.58 ***	(1.18)	6.55 ***	(1.22)	6.44 ***	(1.20)
Age 45 - 54	7.93 ***	(1.51)	7.57 ***	(1.47)	8.21 ***	(1.69)	7.90 ***	(1.67)
Age 55 - 64	10.69 ***	(2.04)	9.99 ***	(2.05)	8.98 ***	(2.01)	8.33 ***	(2.00)
Male*Age 25 - 34	-1.29 **	(0.57)	-1.30 **	(0.57)	-1.49 **	(0.61)	-1.61 ***	(0.60)
Male*Age 35 - 44	-8.00 ***	(1.12)	-7.90 ***	(1.11)	-1.16	(0.74)	-1.20	(0.74)
Male*Age 45 - 54	-14.19 ***	(1.59)	-13.94 ***	(1.55)	-7.00 ***	(1.38)	-6.98 ***	(1.37)
Male*Age 55 - 64	-21.43 ***	(2.08)	-20.83 ***	(1.98)	-12.57 ***	(1.94)	-12.41 ***	(1.92)
Vary by Gender, Age, Race, and Geo								
Non-suicide Mortality Rate (per 100,000)	3.93E-03 ***	(9.2E-04)	3.74E-03 ***	(8.9E-04)	1.85E-03 **	(8.9E-04)	1.76E-03 **	(8.7E-04)
Unemployment Rate	4.28	(4.99)	4.46	(5.08)	20.19 ***	(4.73)	19.47 ***	(4.74)
Male*Unemployment Rate	21.47 ***	(6.61)	21.51 ***	(6.67)	5.94	(8.48)	6.10	(8.44)
Average household size	0.49	(0.62)	0.37	(0.62)	2.18 ***	(0.80)	2.01 **	(0.78)
Share Veteran	26.34 ***	(2.78)	26.31 ***	(2.70)	22.88 ***	(2.96)	23.07 ***	(2.90)
Share married	-25.98 ***	(3.02)	-25.60 ***	(2.76)	-31.65 ***	(4.18)	-30.19 ***	(3.87)
Share widowed	-45.99 ***	(4.65)	-43.51 ***	(4.65)	-43.83 ***	(5.63)	-41.15 ***	(5.44)
Share single /never married	-10.71 ***	(2.79)	-10.50 ***	(2.72)	-13.76 ***	(3.89)	-12.76 ***	(3.69)
Share with < 9th grade ed	-4.05	(3.73)	-4.51	(3.72)	-10.51 **	(4.54)	-10.84 **	(4.54)
Share with < 12th but >= 9th grade ed	1.66	(4.29)	1.97	(4.24)	2.78	(6.19)	2.79	(6.25)
Share with some college	3.78	(2.97)	4.72	(2.95)	-1.10	(2.90)	-0.09	(2.81)
Share with college degree	-4.86	(3.00)	-6.92 ***	(2.63)	-3.15	(3.04)	-5.10 *	(2.83)
Share with MA, prof degree, or PhD	5.32	(4.18)	4.00	(4.25)	-3.06	(4.10)	-4.78	(3.96)
Share Hispanic	0.13	(3.36)	-0.72	(3.40)	-9.69 ***	(2.48)	-10.61 ***	(2.39)
Share of households that live in own home	1.13	(2.64)	2.54	(2.75)	2.92	(2.51)	4.48	(2.53)
Vary by State								
Share of suicides committed by firearm	13.45 ***	(1.92)	12.99 ***	(1.94)	12.61 ***	(2.14)	11.85 ***	(2.02)
Analysis Variables								
Vary by Geo								
Median household income	-1.21E-04 ***	(4.0E-05)	-1.23E-04 ***	(3.8E-05)	-8.90E-05 ***	(0.00)	-9.86E-05 ***	(2.3E-05)
90/10 household income ratio	-0.11	(0.12)			-0.15 *	(0.09)		
90/50 household income ratio			2.50 ***	(0.96)			1.70 ***	(0.46)
50/10 household income ratio			-0.87 **	(0.40)			-0.95 ***	(0.21)
Number of Observations	24,665		24,665		27,750		27,750	
R-Squared	0.4277		0.4285		0.4105		0.4118	

***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively

Table 4: Suicide risk and income inequality—robustness checks, by year

	1990		2000	
	Disaggregated Inequality		Disaggregated Inequality	
Econometric Specification				
State fixed effects	Coefficient	s.e.	Coefficient	s.e.
Median household income	-6.14E-05 *	(3.2E-05)	-6.55E-05 ***	(1.6E-05)
90/50 household income ratio	2.34 ***	(0.79)	1.36 ***	(0.46)
50/10 household income ratio	-0.13	(0.35)	-0.33 **	(0.14)
Number of Observations	24,665		27,750	
R-squared	0.4426		0.4283	
Negative Binomial Poisson MLE				
Median household income	0.78 ***	(6.7E-02)	0.63 ***	(5.0E-02)
90/50 household income ratio	1.49 *	(0.31)	1.35 **	(0.19)
50/10 household income ratio	0.87 **	(0.06)	0.69 ***	(0.05)
Number of Observations	24,635		27,750	
Data and Sample Restrictions				
Extreme values of suicide rate				
Median household income	-1.21E-04 ***	(3.8E-05)	-9.32E-05 ***	(2.2E-05)
90/50 household income ratio	2.89 ***	(1.01)	2.07 ***	(0.50)
50/10 household income ratio	-0.72 *	(0.40)	-0.89 ***	(0.21)
Number of Observations	24,360		27,466	
R-squared	0.5054		0.5014	
Small populations and population weighting				
Median household income	-8.01E-05 *	(4.4E-05)	-6.03E-05 *	(2.8E-05)
90/50 household income ratio	2.62 **	(1.24)	0.31	(0.86)
50/10 household income ratio	-0.80 *	(0.44)	-0.34 *	(0.19)
Number of Observations	17,119		19,519	
R-squared	0.1633		0.1699	
Suicide and accidentdal deaths combined				
Median household income	-3.17E-04 ***	(1.2E-04)	-3.45E-04 ***	(8.2E-05)
90/50 household income ratio	7.49 ***	(2.26)	8.10 ***	(1.58)
50/10 household income ratio	-3.65 ***	(1.11)	-2.93 ***	(0.68)
Number of Observations	24,665		27,750	
R-Squared	0.6462		0.5833	

Table 4: Suicide risk and income inequality—robustness checks, by year (cont.)

Additional explanatory variables (1990 only)

<u>Population density</u>	1.67E-05	(3.0E-05)
Median household income	-1.23E-04 ***	(3.8E-05)
90/50 household income ratio	2.71 ***	(0.96)
50/10 household income ratio	-0.86 **	(0.40)
Number of Observations	24,635	
R-squared	0.4285	
<u>Annual Income growth (1990)</u>	-8.19	(10.29)
Median household income	-1.28E-04 ***	(3.8E-05)
90/50 household income ratio	2.60 ***	(0.95)
50/10 household income ratio	-0.87 **	(0.40)
Number of Observations	24,635	
R-squared	0.4290	
<u>Ten-year income variance (1980-1990)</u>	-6.41E-04	(4.3E-04)
Median household income	-9.74E-05 **	(4.2E-05)
90/50 household income ratio	3.55 ***	(0.94)
50/10 household income ratio	-1.00 ***	(0.38)
Number of Observations	24,635	
R-squared	0.4286	

***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively

All regressions include controls for male, black, other race, 5 year age dummies, male*5 year dummies, non-suicide mortality rate, unemployment rate, male*unemployment rate, average household size, share veteran, married, widowed, single/never married, <9th grade, <12th but >= 9th, some college, college degree, college plus, hispanic, own home, share of suicides committed by firearms in state.

Table 5. Alternative causes of death—robustness checks, by year

	All cause mortality rate							
	1990				2000			
	Overeall Inequality		Disaggregated Inequality		Overall Inequality		Disaggregated Inequality	
Income Variables	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.
Median household income	1.49E-03 **	(7.4E-04)	1.33E-03 **	(6.7E-03)	-1.89E-04	(3.1E-04)	-1.94E-04	(3.2E-04)
90/10 household income ratio	15.55 ***	(2.39)			6.47 ***	(1.21)		
90/50 household income ratio			109.01 ***	(12.95)			39.21 ***	(13.33)
50/10 household income ratio			22.05 ***	(5.87)			13.47 ***	(3.25)
Number of Observations	24,665		24,665		27,750		27,750	
R-Squared	0.9137		0.9139		0.8980		0.8981	

	Deaths from heart failure							
	1990				2000			
	Overeall Inequality		Disaggregated Inequality		Overall Inequality		Disaggregated Inequality	
Income Variables	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.
Median household income	-3.76E-05 **	(1.6E-05)	-3.80E-05 **	(1.5E-05)	-7.08E-07	(1.1E-05)	1.37E-06	(1.3E-05)
90/10 household income ratio	-0.05	(0.06)			0.00	(0.03)		
90/50 household income ratio			-1.41 ***	(0.53)			-0.62	(0.41)
50/10 household income ratio			0.11	(0.21)			0.15	(0.12)
Number of Observations	24,665		24,665		27,750		27,750	
R-Squared	0.2956		0.2961		0.2797		0.2800	

***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively

All regressions include controls for male, black, other race, 5 year age dummies, male*5 year dummies, non-suicide mortality rate, unemployment rate, male*unemployment rate, average household size, share veteran, married, widowed, single/never married, <9th grade, <12th but >= 9th, some college, college degree, college plus, hispanic, own home, and share of suicides committed by firearms in state.

Table 6: Suicide risk and relative income—results based on individual-level data, 1990

Control Variables	OLS		Probit	
Constant and Dummy Variables	Coefficient	s.e.	Coefficient	p-value
Constant	8.324	(21.703)		
Male	14.1754 ***	(0.586)	1.21E-04 ***	0.000
Other Race	-8.7442 ***	(1.034)	-6.62E-05 ***	0.000
Black	-12.3986 ***	(0.744)	-7.41E-05 ***	0.000
Age 25 - 34	-7.0929 ***	(0.941)	-5.96E-05 ***	0.000
Age 35 - 44	-3.6877 ***	(0.708)	-3.35E-05 ***	0.000
Age 45 - 54	0.4710	(0.412)	1.25E-06	0.767
Age 55 - 64	2.2667 ***	(0.327)	1.83E-05 ***	0.000
Male*Age 25 - 34	5.1904 ***	(0.786)	7.48E-05 ***	0.000
Male*Age 35 - 44	2.2427 ***	(0.56)	2.57E-05 ***	0.000
Male*Age 45 - 54	-2.5211 ***	(0.659)	-1.00E-05 ***	0.008
Male*Age 55 - 64	-2.6571 ***	(0.536)	-1.50E-05 ***	0.000
Hispanic	-11.7377 ***	(0.882)	-6.95E-05 ***	0.000
Vary by county-age-race-sex-hispanic-education-marital status				
Share of households that live in own home	-8.64 ***	(2.519)	-6.60E-05 ***	0.000
Unemployment Rate	30.44 ***	(2.724)	1.46E-04 ***	0.000
Share Vietnam veteran	11.03 ***	(2.451)	3.98E-05 ***	0.000
Vary by State				
Share of suicides committed by firearm	21.28 ***	(3.729)	1.57E-04 ***	0.000
Analysis Variables				
Vary by cell definition				
Own income	-7.99 ***	(7.7E-01)	-5.18E-05 ***	0.000
Reference-Group Income (Mean)	7.33 ***	(2.02)	4.69E-05 ***	0.003
Number of Observations	171,192		4,847,659	
(Pseudo) R-Squared	0.0964		0.0562	

***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively

Table A1. Correlation matrices for income variables

	1990			
	Median household income	90/10 ratio of household income	90/50 ratio of household income	50/10 ratio of household income
Median household income	1			
90/10 ratio of household income	-0.4406	1		
90/50 ratio of household income	-0.3533	0.7740	1	
50/10 ratio of household income	-0.4382	0.9596	0.5784	1

	2000			
	Median household income	90/10 ratio of household income	90/50 ratio of household income	50/10 ratio of household income
Median household income	1			
90/10 ratio of household income	-0.3214	1		
90/50 ratio of household income	-0.2096	0.7371	1	
50/10 ratio of household income	-0.3424	0.9393	0.4874	1

Table A2. Regressions by sex and by year, overall inequality

Dependent Variable	1990				2000			
	Male		Female		Male		Female	
Suicide Rate (per 100,000)								
Independent Variables	Coef.	Std. Dev.	Coef.	Std. Dev.	Coef.	Std. Dev.	Coef.	Std. Dev.
Constant and Dummy Variables								
Constant	3.77	(3.04)	51.16 ***	(8.62)	4.80	(3.56)	44.63 ***	(9.28)
Other Race	-0.81	(0.69)	-6.81 ***	(2.26)	0.33	(0.49)	-3.00 *	(1.62)
Black	-4.65 ***	(0.82)	-15.26 ***	(1.81)	-5.83 ***	(0.74)	-16.78 ***	(1.98)
Age 25 - 34	3.41 ***	(0.66)	6.17 ***	(1.80)	3.01 ***	(0.54)	5.13 ***	(1.79)
Age 35 - 44	5.01 ***	(0.82)	2.44	(2.34)	5.03 ***	(0.78)	5.31 **	(2.39)
Age 45 - 54	6.46 ***	(0.97)	-0.06	(3.17)	5.82 ***	(1.03)	3.34	(2.89)
Age 55 - 64	8.99 ***	(1.34)	-0.11	(4.27)	6.21 ***	(1.18)	1.32	(3.33)
Vary by Year, Geo, Sex, Race, and Age								
Non-suicide Mortality Rate (per 100,000)	3.64E-04	(7.4E-04)	4.58E-03 ***	(1.5E-03)	-5.30E-04	(7.2E-04)	2.43E-03 *	(1.3E-03)
Vary by Year, Geo, Sex, and Race								
Unemployment Rate	3.12	(3.28)	31.08 ***	(9.13)	10.89 ***	(3.10)	34.70 ***	(7.30)
Average household size	-0.22	(0.37)	1.20	(1.00)	0.31	(0.47)	3.05 **	(1.19)
Share Veteran	2.81	(4.72)	14.90 ***	(5.18)	5.49	(3.55)	9.02 ***	(3.10)
Share married	-7.64 ***	(2.07)	-51.55 ***	(6.10)	-12.57 ***	(2.91)	-45.22 ***	(8.28)
Share widowed	-16.16 ***	(3.61)	-101.17 ***	(18.81)	-16.72 ***	(4.92)	-73.98 ***	(15.90)
Share single /never married	-2.52	(1.78)	-29.92 ***	(6.47)	-5.82 **	(2.62)	-27.32 ***	(8.04)
Share with < 9th grade ed	-7.18 ***	(1.82)	-15.39 **	(7.42)	-10.08 ***	(2.46)	-18.29 **	(7.67)
Share with < 12th but > 8th grade ed	2.38	(2.36)	-2.30	(7.08)	0.27	(3.54)	-1.67	(8.50)
Share with some college	8.10 ***	(1.82)	3.22	(5.23)	3.52 **	(1.58)	0.96	(4.09)
Share with college degree	-1.23	(1.84)	-15.67 ***	(5.86)	3.19 *	(1.74)	-13.64 **	(5.52)
Share with MA, prof degree, or PhD	9.76 **	(4.46)	11.34 **	(5.77)	-0.94	(3.65)	4.66	(5.77)
Share Hispanic	1.26	(1.92)	1.83	(4.81)	-2.85 **	(1.14)	-13.48 ***	(3.65)
Share of households that live in own home	-1.49	(1.29)	3.06	(4.13)	1.81	(1.43)	2.85	(3.73)
Vary by Year and Geo								
Median household income	-2.18E-04 ***	(6.8E-05)	-3.86E-05 *	(2.2E-05)	-3.40E-05 ***	(1.3E-05)	-1.77E-04 ***	(4.6E-05)
90/10 household income gap	-0.34 *	(0.18)	0.04	(0.07)	-0.02	(0.05)	-0.35 ***	(0.13)
Vary by State								
Share of suicides committed by firearm	6.99 ***	(0.93)	21.38 ***	(3.54)	7.22 ***	(1.17)	19.02 ***	(3.77)
Number of Observations	12,457		12,208		13,773		13,977	
R-Squared	0.1363		0.0860		0.101		0.1822	

*** denotes significance at the 99% level

** denotes significance at the 95% level

* denotes significance at the 90% level

Table A3. Regressions by sex and by year, disaggregate inequality

| | 1990 | | | | 2000 | | | |
| | Male | | Female | | Male | | Female | |
Dependent Variable	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.
Suicide Rate (per 100,000)								
Independent Variables	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.
Constant and Dummy Variables								
Constant	1.85	(2.79)	1.85	(2.79)	2.79	(2.94)	42.51 ***	(9.34)
Other Race	-0.74	(0.70)	-0.74	(0.70)	0.64	(0.48)	-2.27	(1.59)
Black	-4.64 ***	(0.81)	-4.64 ***	(0.81)	-5.74 ***	(0.72)	-16.66 ***	(1.96)
Age 25 - 34	3.47 ***	(0.64)	3.47 ***	(0.64)	3.09 ***	(0.54)	5.14 ***	(1.77)
Age 35 - 44	5.01 ***	(0.81)	5.01 ***	(0.81)	4.92 ***	(0.78)	5.00 **	(2.36)
Age 45 - 54	6.33 ***	(0.97)	6.33 ***	(0.97)	5.57 ***	(1.06)	2.84	(2.89)
Age 55 - 64	8.70 ***	(1.35)	8.70 ***	(1.35)	5.70 ***	(1.25)	0.55	(3.35)
Vary by Year, Geo, Sex, Race, and Age								
Non-suicide Mortality Rate (per 100,000)	4.00E-04	(7.3E-04)	4.00E-04	(7.3E-04)	-4.14E-04	(7.0E-04)	2.30E-03 *	(1.2E-03)
Vary by Year, Geo, Sex, and Race								
Unemployment Rate	3.10	(3.31)	3.10	(3.31)	10.12 ***	(3.04)	34.63 ***	(7.33)
Average household size	-0.28	(0.36)	-0.28	(0.36)	0.18	(0.47)	2.78 **	(1.16)
Share Veteran	3.05	(4.75)	3.05	(4.75)	5.97	(3.64)	9.23 ***	(3.13)
Share married	-7.41 ***	(1.99)	-7.41 ***	(1.99)	-11.58 ***	(2.59)	-42.85 ***	(7.89)
Share widowed	-15.51 ***	(3.57)	-15.51 ***	(3.57)	-15.71 ***	(4.84)	-71.64 ***	(15.65)
Share single /never married	-2.37	(1.72)	-2.37	(1.72)	-5.16 **	(2.43)	-25.92 ***	(7.71)
Share with < 9th grade ed	-7.42 ***	(1.83)	-7.42 ***	(1.83)	-10.31 ***	(2.40)	-18.47 **	(7.68)
Share with < 12th but > 8th grade ed	2.45	(2.37)	2.45	(2.37)	0.04	(3.62)	-1.62	(8.58)
Share with some college	8.30 ***	(1.77)	8.30 ***	(1.77)	3.93 ***	(1.49)	2.69	(4.13)
Share with college degree	-1.98	(1.85)	-1.98	(1.85)	1.85	(1.93)	-16.29 ***	(5.17)
Share with MA, prof degree, or PhD	8.92 *	(4.65)	8.92 *	(4.65)	-2.30	(3.54)	2.29	(5.51)
Share Hispanic	0.97	(1.97)	0.97	(1.97)	-3.37 ***	(1.11)	-14.77 ***	(3.60)
Share of households that live in own home	-0.98	(1.27)	-0.98	(1.27)	2.80 *	(1.50)	5.06	(3.75)
Vary by Year and Geo								
Median household income	-2.18E-04 ***	(6.3E-05)	-3.81E-05 *	(2.2E-05)	-3.95E-05 ***	(1.1E-05)	-1.88E-04 ***	(4.0E-05)
90/50 household income gap	3.55 **	(1.65)	1.16 **	(0.57)	1.35 ***	(0.47)	1.82 **	(0.74)
50/10 household income gap	-1.74 ***	(0.59)	-0.10	(0.23)	-0.39 ***	(0.13)	-1.68 ***	(0.34)
Vary by State								
Share of suicides committed by firearm	6.79 ***	(0.97)	6.79 ***	(0.97)	6.70 ***	(1.15)	18.04 ***	(3.61)
Number of Observations	12,457		12,208		13,773		13,977	
R-Squared	0.1382		0.0864		0.1031		0.1842	

*** denotes significance at the 99% level

** denotes significance at the 95% level

* denotes significance at the 90% level