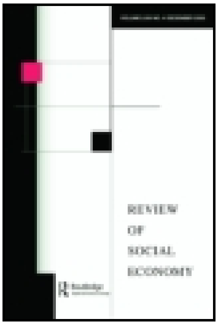

Review of Social Economy
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/rrse20

Local Area Inequality and Worker Well-Being

Michael D. Carr [a]

[a] Department of Economics, University of Massachusetts-Boston, Boston, MA, USA

Version of record first published: 30 Jul 2012

REVIEW OF SOCIAL ECONOMY, 2012, i*First* article

Local Area Inequality and Worker Well-Being

Michael D. Carr
Department of Economics, University of Massachusetts-Boston, Boston, MA, USA

Abstract This paper uses General Social Survey data linked to Census data to investigate the effect of local area income and income inequality on worker well-being. Others have found a robust negative correlation between reference group income and self-reported well-being. However, in many cases the reference group is defined as a large geographic area. This paper adds to the literature in two ways. First, it considers multiple nested geographic reference groups with US data. Second, it explicitly considers income inequality in addition to the level of income. It is found that both income and income inequality are positively associated with well-being at the census tract level, but negatively associated at the county level. Further, the effect of inequality on well-being decreases as income increases at the census tract and county level, while it increases at the state level.

Keywords: subjective well-being, inequality, happiness, satisfaction, comparison income

INTRODUCTION

In the last 20 years, there has been a large increase in research on self-reported well-being. Although this field consists of many areas of study, the core of the literature is still understanding the role of income comparisons in determining subjective well-being. Beginning with the Easterlin Paradox (Easterlin, 1974, 1995, 2001), researchers have documented a seemingly robust negative correlation between reference group income and individual well-being in cross-section regressions, but no correlation in longitudinal data.

Review of Social Economy
ISSN 0034-6764 print/ISSN 1470-1162 online
http://www.tandfonline.com
http://dx.doi.org/10.1080/00346764.2012.707399

Recently, the robustness of the cross-sectional relationship has been questioned. Most notably, the sign and magnitude of the relationship between reference group income and individual well-being is highly sensitive to the size of the geographic area used to define a reference group. In regressions where the reference group is a large geographic area, individual well-being is negatively associated with reference group income (Luttmer, 2005), while for small geographic regions there is a positive association (Barrington-Leigh and Helliwell, 2008; Kingdon and Knight, 2007).

This paper investigates the role of both the level and the distribution of reference group income in shaping subjective well-being, building on recent experimental and theoretical contributions. Two sources of data are used: the General Social Survey and the 2000 Decennial Census Summary File 3. The two datasets are linked, respectively, by census tract, county, and state using a set of geocode identifiers.[1] This nested approach allows the separation of the effect of income inequality from the level of income, and the effect of income and income inequality at each respective geographic level.

Briefly, it is found that census tract income is positively associated with well-being, while county and state income are negatively associated with well-being. Income inequality at the census tract and state levels are both positively associated with well-being, while county inequality is negatively associated with well-being. The results suggest two things. First, the effect of comparison income on well-being is sensitive to how the reference group is defined. Second, the relationship between income and income inequality and immeasurable externalities present at different levels of geographic aggregation matter for well-being in a complex way.

EXISTING LITERATURE

Reference Groups

Studies of interpersonal comparisons are sensitive to the specification of the reference group. A reference group is a group of people that a given individual compares oneself to. The individual can be a member of the group and wants to appear to be a member (conformism), is not a member of the group but wants to a appear to be a member (emulation), is a member of the

1 The geocode identifiers are considered sensitive data and are not publicly available. They can be made available upon request from NORC.

group and wants to appear not to be a member (non-conformism), and is not a member and wants to appear to not be a member (distancing) (Akerlof, 1997).

The primary difficulty for defining reference groups is that they are context specific. Social psychologists argue that individuals tend to compare themselves to others who are similar to them along lines relevant to the context (Burke, 2004). This makes defining the reference group for workers quite complicated. Employed individuals exist within a workplace, where pay comparisons are clearly important (Bowles and Park, 2005; Carr, 2011), within a community, where interpersonal comparisons are also important (Frank, 2005a, 2005b; Luttmer, 2005), and within a more nebulous group that can span both geography and demographics (Frey et al., 2007; Schor, 1998, 2004). The large majority of studies of the effect of reference group income on well-being define the reference group by either geography or demographics. That convention will be followed here for two reasons. First, there is little doubt that we are affected by the decisions and living conditions of our neighbors. Even if one has little direct interaction with one's neighbors, we see their cars, houses, yards, and the like everyday. Second, the groups are easy to define, and mutually exclusive.

This does not undermine the importance of other types of groups. As mentioned above, television and other forms of media put us in contact with the everyday life choices of individuals who are very different both in terms of geographic location and purchasing power. This has been shown to be important for both self-reported well-being (Frey et al., 2007) and consumption decisions (Schor, 1998). These groups, however, are much harder to define, and play a somewhat different role in our lives. It is one thing to demonstrate that individuals who watch more television are less happy and consume more, it is quite another to define a reference group based on television content. If it is the case that causality runs from television watching to consumption—a big if—we can guess about what constitutes the reference group, but we cannot define this group with any certainty.

Put simply, studies that define the reference group by geography or demographics are able to explicitly define a group. Studies that investigate, for example, the effect of television viewing on consumption or happiness infer this relationship by assuming what the reference group must be based on content on television and in other media sources. Without the ability to articulate a well-defined group, there is no way to measure the distribution of outcomes within that group.

Estimation

The literature on the effect of relative income on subjective well-being began largely with the Easterlin Paradox (Easterlin, 1974, 1995, 2001).[2] The Easterlin Paradox is the observation that income is positively correlated with reported well-being at a given point in time, but not over time. Most studies that estimate the effect of relative income on well-being estimate equations like Equation (1).

$$w_{ig} = \alpha + \delta inc_{ig} + \beta \overline{inc}_{-ig} + \gamma X_{ig} \tag{1}$$

where w_{ig} is self-reported well-being of individual i in group g, inc_{ig} is income, $\overline{inc}_{-ig}$ is average income of all individuals aside from i in group g (i.e. reference group income), X_{ig} is a vector of individual controls, Z_g is a vector of reference group controls, a_g is a reference group unobserved effect, and u_{ig} is the individual error term. It is important to note that, depending on the structure of the dataset, a_g is not always explicitly estimated.

The coefficient of interest is β. All else equal, if β is negative, then an increase in reference group income decreases well-being, and vice versa. The large majority of studies that estimate a variant of Equation (1) find $\beta< 0$. This result has been found in multiple countries, with many different reference group specifications, with any number of different control variables, and with varied estimation strategies.[3] Importantly, these studies predominantly use a single reference group, typically defined by either geography or demographics.

This result is simultaneously intuitive and puzzling. On the one hand, evidence from experimental economics and social psychology demonstrates behavior that is consistent with an aversion to both disadvantageous and advantageous inequality.[4] In Equation (1), an increase in $\overline{inc}_{-ig}$ holding ing_{-ig} fixed represents an increase in disadvantageous inequality if $inc_{-ig} < \overline{inc}_{-ig}$, but a decrease in advantageous inequality if the opposite is true. Thus, according to this literature, the sign of β should depend on whether individual income is greater than or less than reference group income. Fehr and Schmidt (1999) argue that individuals dislike disadvantageous inequality

2 There is a longer history of the study of the effect of relative income on utility and consumption, most notably Duesenberry (1949). But, this literature does not pertain to subjective well-being specifically, but instead to broader notions of utility and consumption.

3 See Clark et al. (2008), Frank (2005b), Frey and Stutzer (2002, 2005), Helliwell (2003); Helliwell and Huang (2005), McBride (2001).

4 See Bolton and Ockenfels (2000), Bowles (2008), Camerer (2003), Charness and Rabin (2002), Fehr et al. (1998), Fehr and Schmidt (1999).

more than they dislike advantageous inequality, which would generate $\beta < 0$ in Equation (1) on average.

In addition to this theoretical ambiguity about the sign of β is a more practical issue: reference group income $\left(\overline{inc}_{-ig}\right)$ calculated from surveys should be highly correlated with a whole host of things ranging from health outcomes to social status, that are very difficult to measure independently of the level of income itself, and are also positively correlated with well-being. This conflict can be resolved when reference groups are based on individuals who are observably similar, because grouping in this way implicitly controls for much of the common variation in immeasurable attributes of the individuals. When reference groups are defined by geography, however, it is much more difficult to explain why this negative association is so robust.

It could also be that β varies with the geographic size of the assumed neighborhood. In all of the studies cited above, only one reference group is used, and the level of aggregation is typically quite high. For example Luttmer (2005), using data on the United States, specifies geographic reference groups with approximately 150,000 inhabitants. Helliwell and Huang (2005) improve on this further by using Canadian data at the census tract level, a region with a median population of about 4800. Most studies, however, use reference groups defined as entire states or countries.

A small number of studies have improved on this approach by using multiple levels of geographic reference groups simultaneously. Instead of estimating Equation (1), equations like (2) are estimated.

$$wi_{gh} = \alpha + \delta inc_{igh} + \beta_1 inc_{-ig} + \beta_2 \overline{inc}_{-ig} + \gamma x_{igh} + \theta Z_g + \omega Y_h + a_g + b_h + . \quad (2)$$

where i indexes individuals, g indexes the small reference group which is nested in a larger reference group h, Y_h are characteristics of the larger group h, and b_h is an unobserved effect for h. The other variables are defined as before. Kingdon and Knight (2007) use data on South Africa to estimate a variant of Equation (2). They find that $\beta_1 > 0$ but $\beta_2 < 0$, suggesting that well-being is positively associated with average income of the small group, but negatively associated with average income of the large group. However, it is difficult to assess the general validity of these results, as the sample used is small and not nationally representative. In a working paper, Barrington-Leigh and Helliwell (2008) find similar results using nationally representative Canadian data.

Barrington-Leigh and Helliwell (2008) use responses from several Canadian surveys. The survey data is then merged with Census data at multiple, nested geographic levels. The smallest geographic level used is called a Dissemination Area, and has a median of 480 inhabitants. The largest geographic level used is the Province. In between these two groups, ranging from smallest to largest is the Census Tract, Census Subdivision, and Census metropolitan area. In the context of Equation (2), Barrington-Leigh and Helliwell (2008) have five separate reference group incomes, meaning they estimate five β's. To estimate β_1, the effect of reference group income for the smallest group, fixed effects for all other geographic areas are included. To estimate β_2, the second smallest reference group, fixed effects for the next three larger reference groups are included, and so on. The result is a sequence of regressions that separately estimate the effect of reference group income at each geographic level. Barrington-Leigh and Helliwell (2008) find a positive coefficient on reference group income for the smallest geographic area ($\beta_1 > 0$), and negative coefficients on reference group income at all other geographic levels, respectively.

The main difficulty with this approach is that as the analysis moves from one level of geographic aggregation to another, a number of things are all changing at once. First, reference group likely changes. Second, the number of individuals in the reference group changes. And third, the distribution of outcomes in the reference group changes. These three changes present both conceptual and technical difficulties.

The first is controlled for by either using geographic fixed effects (Barrington-Leigh and Helliwell, 2008), or simultaneously including all levels of geography (Kingdon and Knight, 2007). The second is somewhat more complicated. At higher levels of geographic aggregation, there will be less variation in average reference group income across the reference groups. For example, there is less variation in average income across states in the US than across counties in the US, as all states have high and low income counties. Because of this, it could appear that an individual is less responsive to inequality at higher geographic levels, when in fact this is not correct. The impact of this effect can be handled to a certain degree with estimation strategy. Finally is the effect of the distribution of outcomes, namely that although average incomes converge as the reference group gets larger, inequality will tend to increase.

As mentioned, recent evidence suggests individuals prefer less income inequality. The Fehr and Schmidt (1999) model of inequity aversion, and the data that support it, implies that utility is maximized when there is neither advantageous nor disadvantageous inequality. As long as income inequality and median/mean income are uncorrelated, ignoring the distribution of income will not bias the coefficients on reference group income. But, what the experimental data suggest is that there may be an important independent effect of income inequality.

This could be for a number of reasons. First, holding income fixed, a more unequal distribution of income means more higher income individuals. This could result in a similar set of positive externalities that an increase in income generates. Second, as mentioned, there is evidence to suggest that individuals have a preference against inequality. Third, if mean income and inequality are correlated, it could be that the coefficient on reference group income, particularly for large geographic areas, is actually picking up the effect of inequality. This latter point is important because, if one is to base policy recommendations on this line of research, one might come to a very different conclusion when considering overall inequality versus mean income within a group. Fourth, if income or income inequality is capturing positive externalities, than we might expect interaction effects between the effect of inequality and individual income: individuals at the top of the local income distribution should benefit less from other's income around them than people at the bottom of the income distribution.

METHODOLOGY

Barrington-Leigh and Helliwell (2008) and Kingdon and Knight (2007) use geographic fixed effects combined with multiple levels of nested geographic reference groups to estimate the effect of reference group income on well-being. This approach has several advantages. It is simple to estimate, and fairly straightforward to interpret. The method can also be accomplished with a relatively small number of higher groups, provided there is sufficient variation both within and across groups. The primary limitation of the strategy is that it cannot capture variation across different subgroups within a larger group without explicitly dividing the sample into smaller samples.

The strategy that best accommodates this is multilevel modeling.[5] Multilevel modeling, also referred to as mixed or hierarchical linear modeling, is an estimation technique that explicitly uses the hierarchical structure of the data to estimate both the coefficients and the standard errors. In this sense, it is quite different from the more familiar technique of clustering standard errors, which adjusts the standard errors to account for the fact that standard errors may be invalid when there are multiple observations on the same unit of observation (e.g. tract, county, or state) (Moulton, 1990).

5 See Raudenbush and Bryk (2002) for a good overview of the topic.

Specifically, Equation (3) will be estimated, where i indexes individuals, g indexes groups t, c, and s, and $gini_g$ measures inequality within each respective geographic unit.

$$w_{itcs} = \alpha + \delta inc_{itcs} + \sum\nolimits_g (\beta_g \overline{inc}_{-ig} + \gamma_g gini_g + \rho_g inc_{itcs} gini_g) + a_g \quad (3)$$

Three reference groups are used: the Census Tract, County, and State. Each geographic area is completely nested within the larger one, resulting in a set of mutually exclusive groups. In Equation (3), both the level and the distribution of reference group income are included. Finally, a set of interactions between individual income and reference group inequality is included to capture the fact that the effect of inequality may vary across the income distribution.

This multilevel model—tracts nested in counties nested in states—has one primary drawback. Multilevel models do not use fixed effects, instead they use random effects. However, this potential drawback is outweighed by the fact that, by replacing a fixed intercept for a given level of the data with a random intercept, it is possible to estimate a model with random intercepts for all higher levels of the data simultaneously. This is what gives multilevel modeling the ability to handle interactions between different geographic levels of the dataset.

The expectation is $\beta_g > 0$ for the census tract (the smallest geographic unit), and $\beta_g < 0$ for county and state. The experimental and social psychology literature suggests that $\gamma_g < 0$ for all levels of geographic aggregation, with the caveat that γ_g may be closer to zero at higher levels of aggregation because this represents a greater social distance (Akerlof, 1997). Finally, there are no clear predictions for the interactions (ρ_g), as there is no literature that investigates either how income affects well-being at different levels of inequality nor how inequality affects well-being at different levels of income.

DATA ANALYSIS

Data and Descriptive Statistics

The data come from two sources. Individual data come from the General Social Survey (GSS) for years 1998 to 2008, a nationally representative survey of residents of the United States. The GSS is linked to the Summary File 3 of the 2000 Decennial Census using geocode identifiers. The Census is linked at three geographic levels: census tract, county, and state. The final dataset has 9,087 observations.

There are a number of restrictions posed on the data that results in a substantial number of observations being excluded from the final sample. First, the GSS geocode data at the Census Tract level only go back to 1998. Second, the sample is limited to only those individuals who either report a positive income or report having a part-time or full-time job. Roughly one-third of individuals who report having a job have missing income. These individuals are assigned an income of zero, and included in the analysis, but are identified with a dummy variable in the regressions so as to minimize the bias on the estimated coefficients[6]

Finally, only the Summary File 3 (SF3) of the 2000 Census is used. This is the only publicly available version of the US Census that includes either Census Tract or County identifiers, making it the only feasible dataset. This does pose several limitations. First, the SF3 is actually a dataset of tables, not individual census forms, thus variables like income, race, and education are reported in groups and tabulated by the census bureau. This poses no problems for demographic variables like race and education as they would be tabulated in a similar manner anyway, but it does mean that the gini coefficient is based on categorical income instead of continuous income. The biggest issue that categorical income poses for the gini coefficient is that the top income category is both open ended and relatively low ($200,000 or more), meaning that the gini coefficient will be relatively accurate for lower income neighborhoods, but too small for high income neighborhoods[7] Lastly, the GSS income measure has been converted to constant 2,000 dollars.

Reference group income and income inequality are calculated at the tract, county, and state levels. The level of income is represented by median household income, and inequality is measured by the gini coefficient of household income. From the GSS, individual income is used for two reasons. The primary argument for using individual income is that the GSS reports individual well-being which, given that the sample is limited only to people who have a job, should largely reflect individual income. However, if well-being is affected by observable consumption through a consumption emulation effect, then household income should also play an important role

6 This procedure is analogous to including a dummy variable for individuals with top-coded incomes, a common practice when there are significant numbers of top-coded incomes. An auxiliary regression was run estimating the probability of having missing income based on demographic and human capital variables, the results showed little systematic bias in non-reporting.

7 Unfortunately, there is no good publicly available dataset to test the validity of the inequality measures based on the Census SF3. It is well documented that the Current Population Survey, the other large dataset with State and County identifiers, does not represent the upper portion of the income distribution well (Piketty and Saez, 2003).

in determining individual well-being (Frank, 1985, 2005b; Schor, 1998). For this reason, a dummy variable for whether the individual resides in a two-income household is also included.

Descriptive Statistics

Figure 1 provides a histogram of the answers to the question 'Taken all together, how would you say things are these days? Would you say that you are very happy, pretty happy, or not too happy?' where 1 is 'Not too happy' and 3 is 'Very happy.' Close to 60% of those surveyed fall in the middle category. Because of this, overall well-being will be converted to a dichotomous variable equal to 1 if the individual reports being very happy, and zero otherwise.

Table 1 displays the descriptive statistics for the key variables. Note that average income in the GSS is individual income, which explains why it is less than Census incomes. Mean income levels from the Census data are actually the mean of median incomes for each tract, county, and state, respectively. This, combined with the relatively low top income bracket in the Census compared with the GSS results in average incomes between the GSS and the Census that are somewhat closer together than actual average individual versus household income would normally be.

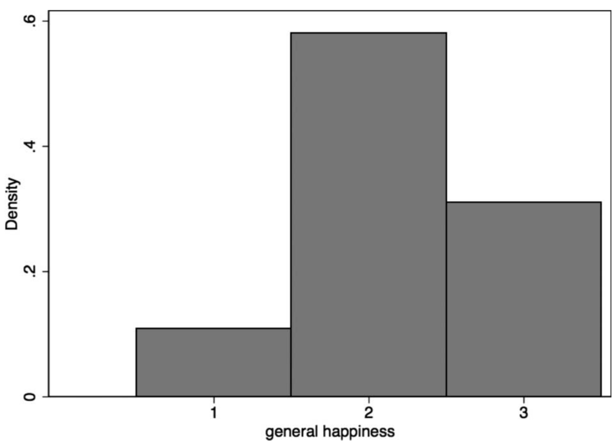

Figure 1: Histogram of General Level of Happiness. *Notes*: Based on author's calculations of GSS data from years 1998 to 2008.

Table 2 breaks down the key variables by overall well-being. Average own income, census tract income, and county income are higher among the high well-being group. But, the difference in means gets smaller at higher levels of geographic aggregation. This pattern is not what is typically seen in

Table 1: Descriptive Statistics.

	Mean	Std. Dev.	*N*
Own income	31368	45393	9087
Tract gini	0.428	0.076	9087
County gini	0.458	0.045	9087
State gini	0.471	0.026	9087
Tract income	45057	18216	9087
County income	43364	10380	9087
State income	42407	5165	9087

Notes: Based on author's calculations of GSS from 1998 to 2008 and 2000 Census data. Own income is individual yearly earnings. Census incomes are household incomes.

Table 2: Descriptive Statistics by Level of Well-being.

	Low	High	T-Stat	Std. err.
Own income	28673	37361	8688.294***	[1022.932]
Tract gini	0.427	0.431	0.004**	[0.002]
County gini	0.459	0.458	−0.001	[0.001]
State gini	0.471	0.47	−0.001	[0.001]
Tract income	44108	47166	3057.350***	[410.877]
County income	43300	43506	206.209	[234.834]
State income	42459	42290	−169.461	[116.845]

Notes: Based on author's calculations of GSS from 1998 to 2008 and 2000 Census data. Own income is individual yearly earnings. Census incomes are household incomes. Low well-being corresponds to categories 1 and 2 of the overall well-being question. High well-being is category 3. N=9,087. Significance levels: *10%, **5%, and ***1%.

regression analysis where higher order reference group income is often negatively correlated with well-being. Census tract inequality is statistically significantly larger for the high well-being group, the opposite of what is implied by most literature on the effect of inequality. Because there are a number of individual and neighborhood characteristics that tend to be correlated with income and potentially income inequality that have nothing to do with income, per se, we turn next to a regression analysis.

Regressions

Table 3 reports the results of four linear probability models. Columns 1 and 3 report the results of an OLS regression. Columns 2 and 4 report the results of a multilevel linear probability model with random effects at the tract, county, and state level. Standard errors in columns 1 and 3 are clustered on the state level because, as suggested by Cameron et al. (2006), when using data with nested clusters it is best to cluster on the highest level cluster.

The coefficients on tract, county, and state income will only be discussed briefly. The results for income are largely consistent with previous work. Well-being is positively associated with income, positively associated with tract income, negatively associated with county income, and negatively associated with state income.

The coefficients on the gini coefficient follow a more complicated pattern: there is a positive coefficient on tract level inequality, a negative coefficient on county level inequality, and a positive coefficient on state level inequality. In columns 1 and 2, the coefficient on tract inequality is statistically significant, implying that a one standard deviation increase in the gini coefficient (16% increase) is associated with a roughly 1-percentage point increase in the probability of reporting high well-being. The coefficients on state and county inequality, though not statistically significant, are considerably larger than the coefficient on tract inequality. Given the relatively small number of counties (317) and states (46) observed in the dataset, it seems reasonable to conclude that both county and state inequality would be statistically significant if there were a way to increase the number of observed counties, in particular. Note also that only the coefficient on state inequality shows any marked change when the random effects are included.

Columns 3 and 4 report the results of the linear probability and multilevel models with interaction terms for own income and tract inequality, county inequality, and state inequality, respectively. There are two ways to interpret the results. When the interactions are included, neither the direct effect nor

Table 3: Results Linear Probability Model for Overall Well-being.

	OLS[a]	Multilevel[b]	OLS[a]	Multilevel[b]
Log own income	0.0259***	0.0259***	0.0183	0.0197
	[0.0054]	[0.0054]	[0.0205]	[0.0205]
Tract gini	0.1275*	0.1302*	0.2037	0.2042
	[0.0722]	[0.0724]	[0.1414]	[0.1414]
County gini	−0.2256	−0.2253	0.0102	0.0133
	[0.1405]	[0.1417]	[0.2698]	[0.2705]
State gini	0.2479	0.3316	−0.1786	−0.0762
	[0.2185]	[0.2554]	[0.4210]	[0.4434]
Log tract income	0.0427**	0.0422**	0.0428**	0.0423**
	[0.0205]	[0.0206]	[0.0205]	[0.0207]
Log county income	−0.0288	−0.0251	−0.0287	−0.0249
	[0.0280]	[0.0282]	[0.0280]	[0.0282]
Log state income	−0.0978**	−0.1006*	−0.1006**	−0.1034**
	[0.0480]	[0.0525]	[0.0480]	[0.0525]
Own inc. x tract gini			−0.0095	−0.0092
			[0.0151]	[0.0151]
Own inc. × county gini			−0.0304	−0.0308
			[0.0295]	[0.0295]
Own inc. × state gini			0.0551	0.0523
			[0.0465]	[0.0466]
N	9087	9087	9087	9087
Random intercepts	No	Yes	No	Yes

Notes: Data are from GSS and Census. Additional controls include marital status, gender, highest degree obtained, age, age squared, race/ethnicity, tract unemployment rate, share of college graduates in the tract, share of tract population that is white, share of tract population that is African American, whether there are other earners in the family, a dummy for having a job but reporting no income, and a set of year dummies. Significance levels: *10%, **5%, and ***1%. [a]Linear probability model with standard errors clustered on the state. [b]Multilevel linear probability model with random intercepts at the census tract, county, and state level.

the interacted effect of own income or any of the inequality measures are statistically significant. However, comparing across similar estimation techniques (i.e. column 1 versus column 3 and column 2 versus column 4), it is clear that including the interaction terms has a large impact on the results. The coefficient on the tract gini increases by about 60%, while the coefficients on county and state inequality, respectively, switch signs resulting in a change in magnitude of more than 100%. In column 4, a one standard deviation increase in the direct effect of tract inequality is now associated with a 1.6-percentage point increase in the probability of reporting high well-being, however given the negative interaction term, this effect decreases as income increases.

In sum, the average individual experiences positive externalities from living in high income neighborhoods, but negative externalities from living in a high income county or state, all else equal. Similarly, the average individual receives a positive benefit from living in a high inequality neighborhood, a negative externality from living in a high inequality county, and a positive externality from a high inequality state overall. But, the estimated impact of inequality depends crucially on whether own income interactions are included.

Interactions

The regressions contain three interactions: log household income by tract inequality, log household income by county inequality, and log household income by state inequality. The logic behind the choice of interactions comes from existing experimental research on the effects of inequality, which suggests that the effect of inequality may vary with one's position in the income distribution[8] In order to understand the full impact of inequality on well-being, both the direct and interacted effects must be taken into account.

Interactions between two continuous variables are complex because the estimated coefficient changes depending on the point that the interaction is evaluated at. This means that values of one of the interacted variables must be chosen in order to evaluate the effect of the other interacted variable. Second, the statistical significance of the interaction depends on where the interaction is evaluated. Thus, although the coefficient on the interacted variables may be statistically insignificant in the regression—as they are in this case—it is possible that the interactions will be statistically significant when evaluated at other points.

The most common method for assessing interactions to evaluate the coefficients (the direct and interacted effects) at 'meaningful' levels of one of

8 See Camerer (2003) for a review of the literature.

the interacted variables. Then, one would typically graph the results in well-being/income space. The limitation of this approach is that it is impossible to graph the estimated coefficient for every level of income. Additionally, because well-being is a dichotomous variable, graphs in income/well-being space are not particularly useful. Here, a second method is used. First, the marginal effect of inequality on well-being is estimated at every observed level of individual income. Second, all of the marginal effects are plotted. The resulting graph is a plot of the first derivative of well-being with respect to each respective measure of inequality, computed at each level of income. All other variables are held constant at their respective means.

Figures 2–4 show the results for census tract, county, and state income. The solid portion of the line is statistically significant at the 10% level or lower, the dashed portion is statistically insignificant at the 10% level. The dotted gray line represents plus/minus one standard error of the estimated coefficient. The magnitude of the coefficients in each figure is based on the results of column 4 from Table 3.

Figure 2 shows the aggregate effect of census tract inequality. There is a clear downward trend in the marginal effect as income increases. At the lowest level of income, the association between inequality and well-being is about 0.15. This declines steadily to about 0.1 as income increases.

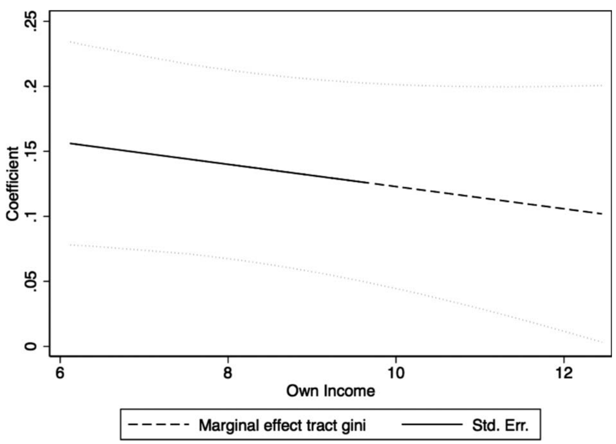

Figure 2: Interaction between Own Income and Census Tract Inequality. *Notes*: Based on author's calculations of GSS data. Interactions estimated at each level of log income, and the mean of all other variables. Based on linear combination of the coefficients on own income and own income × census tract inequality from column 2 of Table 3. Solid line represents statistical significance at the 10% level.

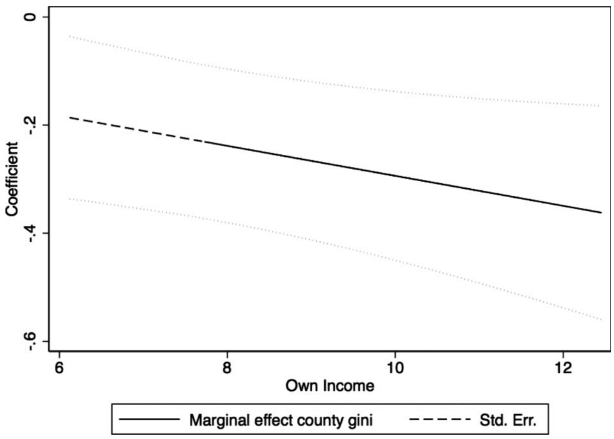

Figure 3: Interaction Between Own Income and County Inequality. *Notes*: Based on author's calculations of GSS data. Interactions estimated at each level of log income, and the mean of all other variables. Based on linear combination of the coefficients on own income and own income × county inequality from column 2 of Table 3. Solid line represents statistical significance at the 10% level.

For county inequality in Figure 3, the change in magnitude is equally as large. For low income counties, the marginal effect is approximately −0.2. This decreases steadily with own income to −0.38. Note also that unlike tract inequality which is statistically insignificant for high income individuals, county inequality is statistically significant for high income individuals.

The aggregate effect of state inequality is shown in Figure 4. Once again, there is variation across levels of income. At low levels of income, the marginal effect is 0.22 and statistically insignificant. This increases steadily to about 0.55 and becomes statistically significant. The effect of tract inequality decreases (i.e. moves closer to zero) as income increases, while the effect of county inequality and state inequality increases (i.e. moves further away from zero) as income increases, though county inequality becomes larger negative as income increases.

SUMMARIZING THE RESULTS

In summary, individual income is positively associated with well-being, though the effect is not significant when the interactions are included. Both tract income and tract inequality are positively associated with well-being, though the association between inequality and well-being decreases as income increases. Both county income and inequality are negatively

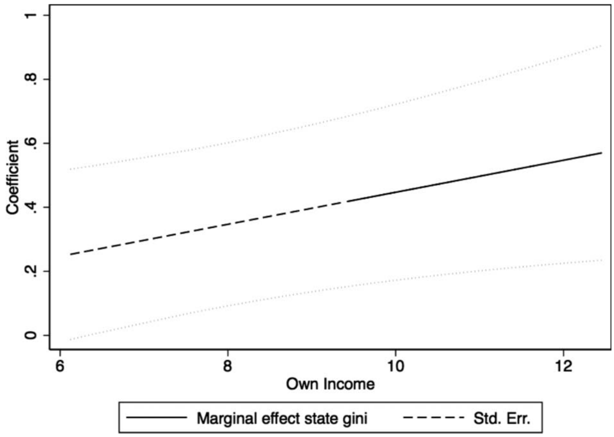

Figure 4: Interaction Between Own Income and State. *Notes*: Based on author's calculations of GSS data. Interactions estimated at each level of log income, and the mean of all other variables. Based on linear combination of the coefficients on own income and own income × state inequality from column 2 of Table 3. Solid line represents statistical significance at the 10% level.

associated with well-being. Further, the negative association between county inequality and well-being becomes stronger as income increases. Finally, state income is negatively associated with well-being, while state inequality is positively associated with well-being, and the association gets stronger as income increases.

Providing a sense of the magnitude of these effects is difficult because, to the author's knowledge, this is the first study to explicitly estimate the relationship between inequality and subjective well-being. Second, the magnitude of the relationship between inequality and well-being varies with income. Third, income and income inequality are measured on different scales, thus one cannot directly compare the effect of e.g. tract income versus tract inequality on well-being. Informally, based on the multilevel regression without interactions, a 1% increase in tract income increases the probability of reporting a high level of well being by 4.3%. A 1% increase in tract inequality, on the other hand, increases the probability by 12.8%. Both effects are large when compared to a 1% increase in own income, which is associated with a 2.6% increase in the probability of reporting high well-being.

Further, as can be seen from the graphs of the interactions in Figures 1–3, the total effect of inequality on well-being varies in magnitude by as much as

250% across the income distribution. The effect of tract inequality on well-being is 50% stronger for low-incomes (0.15) than high incomes (0.1). The effect of county inequality is about 85% larger for high incomes (-0.2) than low-incomes (-0.375). And, the effect of state inequality on well-being is 250% larger for high incomes (0.55) than low incomes (0.22).

DISCUSSION

The primary goal of this paper is to investigate whether inequality is related to individual well-being, and whether this relationship varies with the level of income. The results suggest that the answer to both of these questions is yes. Understanding why this is the case requires substantially more research. Some preliminary suggestions that are consistent with the results will be offered here, but these should not be interpreted as conclusive in any way.

One can think of the effect of inequality in two different ways. Inequality can have a direct effect because of preferences for or against unequal outcomes, as suggested by the experimental and a portion of the theoretical literature. Or, inequality can function as a summary measure of the context within which an individual lives. Given that the experimental literature unambiguously predicts that inequality should be negatively correlated well-being, it would seem that any plausible explanation of the results found here must include an important role for inequality as a summary measure of context.

The biggest puzzle is the result that lower income individuals prefer high inequality neighborhoods, while high income individuals prefer high inequality states. This puzzle can be partly resolved by considering the fact that living in a low inequality neighborhood means living around either entirely low-income or entirely high-income individuals. Given the general shape of the income distribution in the United States, both higher levels of income and inequality are associated with more high income people in a neighborhood. But, because high-income individuals are less dependent upon the neighborhood for amenity provision, there is an asymmetry in the benefits of inequality: lower income individuals benefit from local inequality while higher income individuals benefit from wider inequality. Further, a high-income individual may receive some benefit from living in a low inequality, high income neighborhood, but this benefit will be swamped by the benefit a lower income individual receives from living in a high inequality neighborhood.

Admittedly, fitting the large negative effect of county inequality into the amenity provision hypothesis is difficult. However, there are many important amenities that are provided at the city level and the state level but not the county level. These amenities may not take the same form. Census tracts, and the towns they are in, are closely associated with the quality of public schools, trash removal, snow removal, and other tangible amenities. But, neighborhoods also play an important role in overall quality of life in countless other ways. States play an important role in supporting labor markets, highway road maintenance, and public primary and secondary education that are related as much to the level of state income as they are to the density of high income individuals. Thus, county inequality, holding neighborhood and state inequality fixed, likely captures the negative externalities discussed earlier of living in a higher inequality area.

This line of reasoning is speculative; there is little systematic research into this question using nationally representative surveys. What does exist strongly supports neighborhood choice driven by local amenities, and human capital and demographic characteristics along the lines discussed here (Ioannides and Zabel, 2008; Zanella and Ioannides, 2007). It is also entirely possible that each dimension of inequality matters for very different reasons. What the results do clearly suggest is that, as Frank and Levine (2007) and Frank (2005) argues, some redistribution may be welfare improving. The results imply that, especially at the tract level, the benefits associated with living around high income people play a critical role in the effect of inequality, and it is likely the case that any redistribution that brings up the bottom portion of the income distribution without significantly impacting the top of the income distribution will result in an aggregate increase in welfare.

ACKNOWLEDGEMENTS

The author would like to thank Arjun Jayadev, discussants from the 2010 Eastern Economics Associations meetings, and the anonymous referees. All errors are my own.

REFERENCES

Akerlof, G.A. (1997) "Social Distance and Social Decisions," *Econometrica* 65(5): 1005–1027.

Barrington-Leigh, C. P. and Helliwell, J. F. (2008) "Empathy and Emulation: Life Satisfaction and the Urban Geography of Comparison Groups," *NBER Working Paper Series 14593*.

Bolton, G. E. and Ockenfels, A. (2000) "ERC: A Theory of Equity, Recriprocity and Competition," *American Economic Review* 90(1): 166–193.

Bowles, S. (2008) ''Policies Designed for Self-Interested Citizens May Undermine 'The Moral Sentiments': Evidence from Economic Experiments,'' *Science* 20(320): 1605–1609.

Bowles, S. and Park, Y. (2005) ''Emulation, Inequality, and Work Hours: Was Thorsten Veblen Right?,'' *Economic Journal* 115(507): F397–412.

Burke, P. J. (2004) ''Identities and Social Structure: The 2003 Cooley-Mead Award Address,'' *Social Psychology Quarterly* 67(1): 5–15.

Cameron, A. C., Gelbach, J. B. and Miller, D. L. (2006) ''Robust Inference with Multi-way Clustering,'' *NBER Technical Working Paper No.327*.

Camerer, C. (2003) *Behavioral Game Theory: Experiments in Strategic Interaction,* Princeton: Princeton University Press.

Carr, M. D. (2011) ''Work hours and wage inequality: Evidence from the 2004 WERS,'' *Journal of Socio-economics* 40(4): 417–427.

Charness, G. and Rabin, M. (2002) ''Understanding Social Preferences with Simple Tests,'' *Quarterly Journal of Economics* 117(3): 817–869.

Clark, A. E., Frijters, P. and Shields, M. A. (2008) ''Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles,'' *Journal of Economic Literature* 46(1): 95–144.

Duesenberry, J. S. (1949) *Income, Saving, and the Theory of Consumer Behavior,* Cambridge: Cambridge University Press.

Easterlin, R. A. (2001) ''Income and Happiness: Towards a Unified Theory,'' *Economic Journal* 111(473): 465–484.

Easterlin, R. A. (1974) ''Does Economic Growth Improve the Human Lot? Some Empirical Evidence,'' in R. David and M. Reder (eds) *Nations and Households in Economic Growth: Essays in Honor of Moses Abramovitz*. New York: Academic Press, pp. 89–125.

Easterlin, R. (1995) ''Will Raising the Incomes of All Increase the Happiness of All?'' *Journal of Economic Behavior and Organization* 27(1): 35–48.

Fehr, E., Kirchler, E., Weichbold, A., and Gachter, S. (1998) ''When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets'' *Journal of Labor Economics* 16(2): 324–351.

Fehr, E. and Schmidt, K. M. (1999) ''A Theory of Fairness, Competition, and Cooperation,'' *Quarterly Journal of Economics* 114(3): 817–868.

Frank, R. H. (1985) ''The Demand for Unobservable and Other Nonpositional Goods,'' *American Economic Review* 75(1): 101–116.

Frank, R. H. (2005a) ''Does Absolute Income Matter?'' in Luigino Bruni and Pier Luigi Porta (eds) *Economics and Happiness*. Oxford: Oxford University Press, pp. 65–90.

Frank, R. H. (2005b) ''Positional Externalities Cause Large and Preventable Welfare Losses,'' *The American Economic Review* 95(2): 137–141.

Frey, B. S., Benesch, C. and Stutzer, A. (2007) ''Does Watching TV Make us Happy?,'' *Journal of Economic Psychology* 28(3): 283–313.

Frey, B. S. and Stutzer, A. (2002) *Happiness in Economics,* Princeton: Princeton University Press.

Frey, B. S. and Stutzer, A. (2005) ''Testing Theories of Happiness,'' in L. Bruni and P. L. Porta (eds) *Economics and Happiness,* Oxford: Oxford University Press, pp. 116–146.

Helliwell, J. F. (2003) "How's Life? Combining Individual and National Variables to Explain Subjective Well-Being," *Economic Modeling* 20(3): 331–360.

Helliwell, J. F. and Huang, H. (2005) "How's the Job? Well-Being and Social Capital in the Workplace", *NBER Working Paper Series 11759*.

Ioannides, Y. M. and Zabel, J. E. (2008) "Interactions, Neighborhood Selection, and Housing Demand," *Journal of Urban Economics* 63(1): 229–252.

Kingdon, G. G. and Knight, J. (2007) "Community, Comparisons and Subjective Well-Being in a Divided Society," *Journal of Economic Behavior and Organization* 64(1): 69–90.

Luttmer, E. F. P. (2005) "Neighbors as Negatives: Relative Earnings and Well-Being," *The Quarterly Journal of Economics* 120(3): 963–1002.

McBride, M. (2001) "Relative-Income Effects on Subjective Well-Being in the Cross-Section," *Journal of Economic Behavior and Organization* 45(1): 251–278.

Moulton, B. R. (1990) "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *The Review of Economics and Statistics* 72 (2): 334–338.

Piketty, T. and Saez, E. (2003) "Income Inequality in the United States, 1913–1998," *Quarterly Journal of Economics* 118 (1): 1–39.

Raudenbush, S. W. and Bryk, A. S. (2002) *Hierarchical Linear Models: Applications and Data Analysis Methods,* Thousand Oaks: Sage Publications.

Schor, J. B. (1998) *The Overspent American: Upscaling, Downshifting, and the New Consumer,* New York: Basic Books.

Schor, J. B. (2004) *Born to Buy: The Commercialized Child and the New Consumer Culture,* New York: Scribner.

Zanella, G. and Ioannides, Y. M. (2007) "Searching for the Best Neighborhood: Mobility and Social Interactions," *University of Sienna Working Paper*.

NOTES ON CONTRIBUTORS

Michael D. Carr is an Assistant Professor of Economics at the University of Massachusetts-Boston. His work, both experimental and non-experimental, focuses on the causes and consequences of income, wealth, and power inequality. He received his PhD in Economics from the University of Massachusetts, Amherst in 2009.