Economic Inequality and Class Consciousness in the United States*

Frederick Solt
frederick-solt@uiowa.edu

Yue Hu
yue-hu-1@uiowa.edu

Kevan Hudson
kevan-hudson@uiowa.edu

Jungmin Song
jungmin-song@uiowa.edu

Dong 'Erico' Yu
dong-yu@uiowa.edu

July 17, 2016

Abstract

*Revision history and the materials needed to reproduce the analyses of this paper can be found on Github here.

One of the most important questions underlying recent research on economic inequality and democracy is whether inequality in democratic contexts is self-correcting or instead self-reinforcing. Rational choice arguments have long maintained that, where economic inequality is higher, the benefit of more redistributive policies to the median voter and to those with below-median incomes will be greater, leading them to demand and achieve the adoption of higher taxes and more government spending to ameliorate unequal conditions between those with higher and lower incomes (see, e.g., Meltzer and Richard 1981).

[unequal democracy / relative power stuff]

A prominent recent study, Newman, Johnston, and Lown (2015*a*, hereafter NJL), argues in favor of the former, more optimistic view. It concludes that simply being exposed to high levels of local income inequality prompts those with lower incomes to become more likely to view the United States as divided into haves and have-nots and to see themselves as among the have-nots, that is, to become more likely to achieve a class consciousness vital to contesting the fairness of the economic system and demanding more redistribution.[1]

NJL reaches this conclusion on the basis of analyses of the September 2006 Pew News Interest Index Survey, which it describes as "containing a unique set of questions tapping perceptions of economic hierarchy and inequality and respondents' perception of their own position within such a hierarchy" (p.336). These questions, however, were not at all unique

[1]NJL also purports to show that lower-income people are more likely to reject, and higher-income people are more likely to accept, the meritocratic ideal that hard work leads to success when living in contexts of greater local income inequality. Solt et al. (2016), however, documents how this conclusion is not in fact supported by the results presented in NJL but instead is based on a crucial misinterpretation of a multiplicative interaction term (see Brambor, Clark, and Golder 2006). In an independent replication that brings more and better data to the question, Solt et al. (2016) finds that those with lower incomes are actually *less* likely to reject meritocracy where income inequality is greater.

to the 2006 survey: they were in fact included in no fewer than six Pew surveys during the period examined in the study, 2005 to 2009. Given the longstanding admonition to maximize the number of observations used to test a theory's implications (see, e.g., King, Keohane, and Verba 1994), this is a surprising oversight.[2]

To examine whether the choice of data affected the results presented in NJL, we first reproduce Table 2 from that article using the 2006 Pew survey alone, which presents results regarding views of the United States as divided into 'haves' and 'have-nots.' The NJL reproducibility materials note that the article's authors are themselves unable to reproduce the results exactly (Newman, Johnston, and Lown 2015*b*), though we confirm that the differences between the published results and those obtained are indeed quite small.[3] Next, we replicate this analysis using all six of the available Pew surveys pooled together.[4] In both cases, we follow NJL and examine only views among white respondents. Figure 1 displays

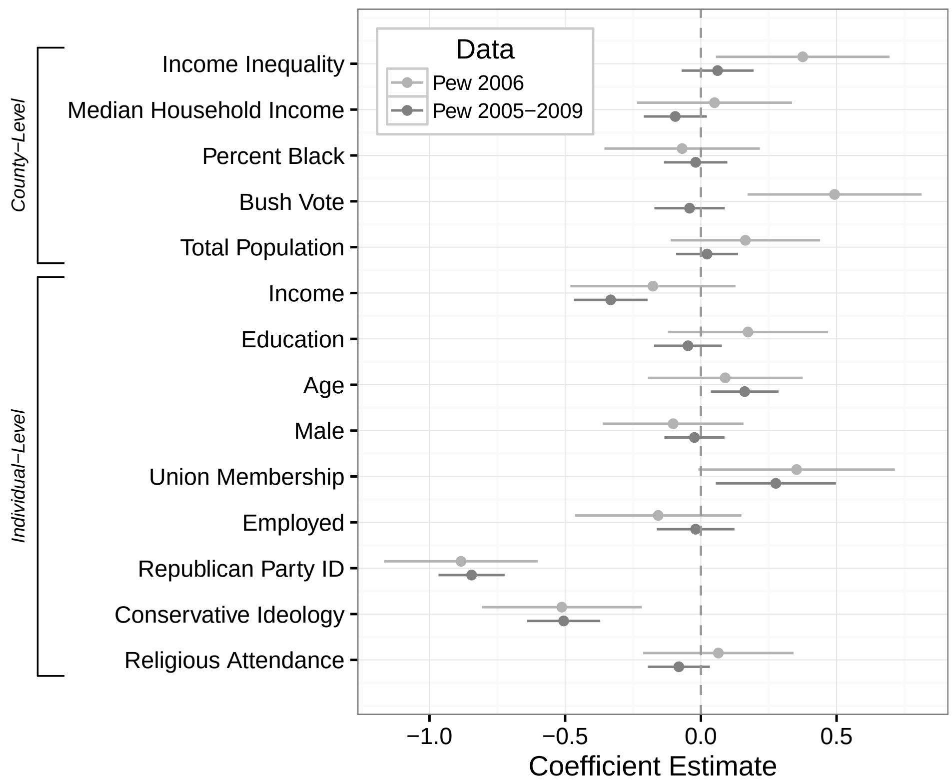
[2] It is perhaps even more surprising in light of the fact that one of these additional surveys, Pew's April 2009 Values Survey, was included in NJL's analyses of meritocratic beliefs even though it did *not* ask the same item as other surveys pooled in those analyses (see NJL, p.331; Solt et al. 2016, 8).

[3] In reproducing the results directly from the original 2006 survey, we discovered two additional minor issues. First, although NJL describes its model as including a control for unemployment (p.331), the survey only includes an item regarding *employment* status; that is, the unemployed cannot be distinguished from students, retirees, and others not in the workforce in these data. We therefore simply more accurately label this variable as "Employed." Second, missing data in the survey appear to have been singly imputed in the reproducibility materials using an undocumented procedure. Following the advice that *multiple* imputation is the best way to preserve observations with missing data without understating the uncertainty due to missing values (see, e.g., Rubin 1987), we use the `R` package `mi` to deal with this issue (Su et al. 2011). Neither of these changes yield substantial differences from the results reported in NJL.

[4] We note that items for three control variables were not asked in all six surveys: union membership (omitted from the July 2007 and October 2008 surveys), employment (July 2007), and church attendance (October 2008). To deal with this issue, we pooled the surveys before multiply imputing the missing data (see Gelman, King, and Liu 1998).

the results as a dot-and-whisker plot (see Kastellec and Leoni 2007; Solt and Hu 2015*a*), with the dots representing the estimated change in the logged odds of the dependent variable for a change of two standard deviations in the independent variable and the whiskers representing the 95% confidence intervals of these estimates.

Figure 1: Local Inequality and the Perception of America as Divided into 'Haves' and 'Have-Nots': Results Using All Available Data



Notes: The dots represent the estimated change in the logged odds of believing the United States to be divided into 'haves' and 'have-nots' for a change of two standard deviations in the independent variable; the whiskers represent the 95% confidence intervals of these estimates. Results from replications of the model presented in Table 2 of Newman, Johnston, and Lown (2015*a*) on the 2006 Pew survey analyzed in that article and on pooled data from the six Pew surveys that included the same item and were conducted in the time period the article examines. The statistically significant result for county income inequality in the 2006 survey presented in that article is not evident when all of the available data are examined.

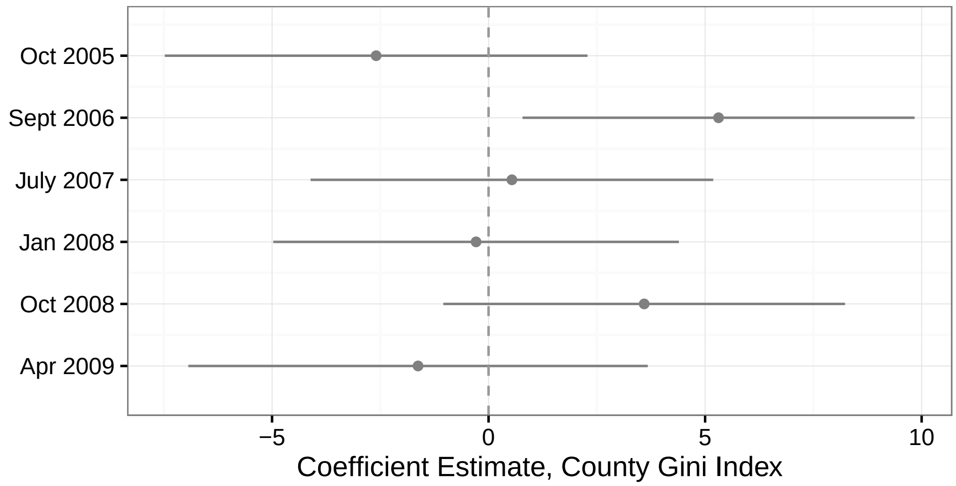
The upper, lighter lines depict the results obtained using only the 2006 survey as in

NJL; the lower, darker lines are those using all of the available surveys. There are many similarities: in both sets of results, most of the estimates do not reach statistical significance, but Republicans and conservatives are less likely to view the United States as divided and union members are more so. The differences, however, are telling. The counterintuitive finding in the 2006 data that people in counties where George W. Bush won a larger share of the vote in 2004 were more likely to see a divide between haves and have-nots evaporates when all of the available surveys are examined, and the surprising null result for income gives way to the expected strongly negative relationship. Most importantly, while the estimate for the context of local income inequality in the 2006 data is positive and statistically significant, this evidence that "income inequality and relative economic comparisons will become more salient among citizens residing in high- than low-inequality contexts" (NJL, p.337) disappears when all of the available surveys are included in the sample.

Moreover, consider Figure 2. It is a "secret weapon" plot (see Gelman 2008, 198) that displays the result for income inequality when the NJL model is fit to each of the six available Pew surveys separately. The coefficient for income inequality is *only* estimated to be positive and statistically significant in the 2006 survey employed in NJL. Not one of the other five datasets yields a statistically significant coefficient, and in three of them the point estimate is actually negative. "We should be willing", King, Keohane, and Verba (1994, 31) advise, "to take whatever information we can acquire so long as it helps us learn about the veracity of our theory." In neglecting to examine more than a single survey, NJL reaches a conclusion that is not supported by the available evidence.

Are lower-income citizens living in localities with more income inequality more likely to see themselves as have-nots? The NJL analysis of this question, presented in that article's

Figure 2: Local Inequality and the Perception of America as Divided into 'Haves' and 'Have-Nots': Results Using Each Available Dataset

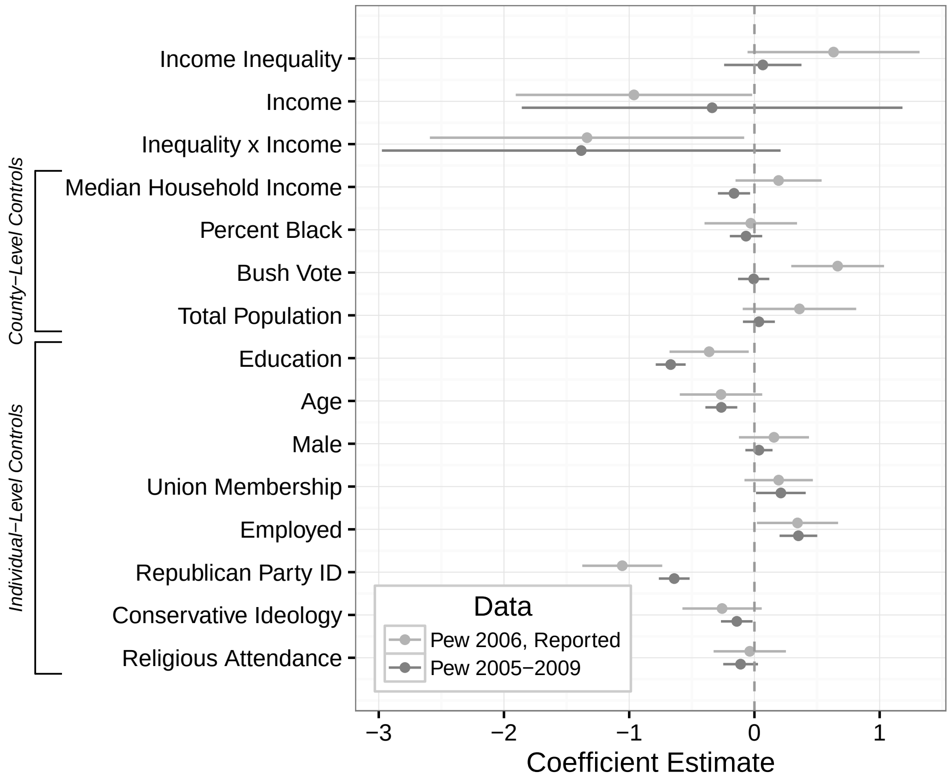


Notes: The dots represent the estimated change in the logged odds of believing the United States to be divided into 'haves' and 'have-nots' for a change of two standard deviations in county income inequality; whiskers represent 95% confidence intervals. Results for county income inequality from replications of the model presented in NJL Table 2 on data from each of six available surveys conducted in the in the time period examined in that article. Of the six surveys, the only one that yields a statistically significant result is the 2006 survey presented in the article.

Table 3, is not reproducible. The 2006 survey comprises just 1067 white respondents, but those respondents live in 661 counties and the NJL model specifies two random effects per county, yielding over 1300 parameters (see NJL, p.336). This is the classic 'small-n' problem writ large: it is impossible to estimate the model without additional information (see, e.g., King, Keohane, and Verba 1994). Figure 3 therefore displays the reported results for reference, alongside the results obtained by estimating the model using all six Pew surveys that asked respondents if they identified themselves as among the have-nots.

Note that the NJL model includes a multiplicative interaction term between county-level income inequality and individual income. For this reason, the coefficients of these variables are not interpretable directly; instead, to understand how these variables relate to self-identification as a have-not, their conditional effects must be calculated and plotted (see,

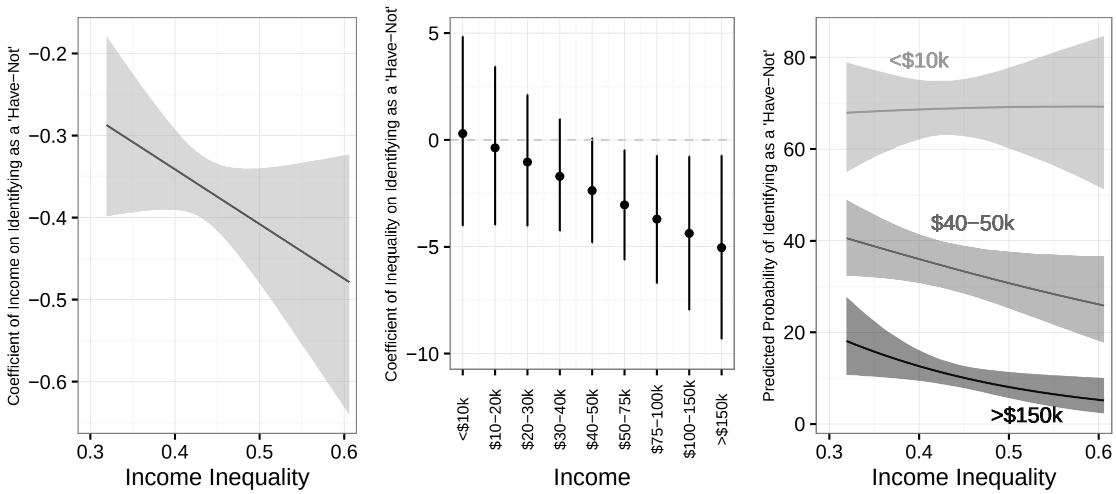
Figure 3: Local Inequality and Self-Identification as a ‘Have-Not’: Results Using All Available Data



Notes: The dots represent the estimated change in the logged odds of self-identifying as a ‘have-not’ for a change of two standard deviations in the independent variable; whiskers represent 90% confidence intervals corresponding to the one-tailed tests applied in NJL. Results obtained by replicating the analysis from NJL Table 3 on pooled data from the six Pew surveys that included the same item conducted in the time period examined in that article. The published NJL analysis, purportedly of the 2006 Pew survey, is not reproducible because it includes more parameters than observations; the reported results are depicted here.

e.g., Brambor, Clark, and Golder 2006). Using the `R` package `interplot`, we present these conditional effects in Figure 4.

Figure 4: Conditional Effects of Income and Local Inequality on Self-Identification as a 'Have-Not'



Notes: The dots and solid lines represent the estimated change in the logged odds of the dependent variable for a change of two standard deviations in the independent variable; whiskers and shaded regions represent 95% confidence intervals. Income is estimated to have a negative effect on identifying as a have-not that is strong, statistically significant, and larger in magnitude as local income inequality increases. There is no support for the conclusion reached in NJL that lower income people are more likely to identify as have-nots when they live in contexts of greater income inequality; there is only support for the proposition that those with higher incomes are less likely to do so where inequality is greater.

References

Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. “Understanding Interaction Models: Improving Empirical Analyses.” *Political Analysis* 14(1):63–82.

Gelman, Andrew. 2008. *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do.* Princeton: Princeton University Press.

Gelman, Andrew, Gary King, and Chuanhai Liu. 1998. “Not Asked and Not Answered: Multiple Imputation for Multiple Surveys.” *Journal of the American Statistical Association* 93(443):846–857.

Kastellec, Jonathan P., and Eduardo L. Leoni. 2007. “Using Graphs Instead of Tables in Political Science.” *Perspectives on Politics* 5(4):755–771.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton: Princeton University Press.

Meltzer, Allan H., and Scott F. Richard. 1981. “A Rational Theory of the Size of Government.” *Journal of Political Economy* 89(5):914–927.

Newman, Benjamin J., Christopher D. Johnston, and Patrick L. Lown. 2015*a*. “False Consciousness or Class Awareness? Local Income Inequality, Personal Economic Position, and Belief in American Meritocracy.” *American Journal of Political Science* 59(2):326–340.

Newman, Benjamin J., Christopher D. Johnston, and Patrick L. Lown. 2015*b*. “Replication data for: False Consciousness or Class Awareness? Local Income Inequality, Personal Economic Position, and Belief in American Meritocracy.” http://dx.doi.org/10.7910/DVN/26584, Harvard Dataverse, V2.

Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys.* New York: J. Wiley & Sons.

Solt, Frederick, and Yue Hu. 2015*a*. “dotwhisker: Dot-and-Whisker Plots of Regression Results.” Available at the Comprehensive R Archive Network (CRAN).

Solt, Frederick, and Yue Hu. 2015*b*. “interplot: Plot the Effects of Variables in Interaction Terms.” Available at the Comprehensive R Archive Network (CRAN).

Solt, Frederick, Yue Hu, Kevan Hudson, Jungmin Song, and Dong Yu. 2016. “Economic Inequality and Belief in Meritocracy in the United States.” https://github.com/fsolt/meritocracy/blob/master/paper/merit.pdf.

Su, Yu-Sung, Andrew Gelman, Jennifer Hill, and Masanao Yajima. 2011. “Multiple Imputation with Diagnostics (mi) in R: Opening Windows into the Black Box.” *Journal of Statistical Software* 45(2):1–31.