THE SOCIAL DYNAMICS OF ECONOMIC POLARIZATION:

EXPLORING THE LIFE COURSE PROBABILITIES OF

TOP-LEVEL INCOME ATTAINMENT

Thomas A. Hirschl
Department of Development Sociology
Cornell University
Ithaca, New York 14853
(607) 255-1688
tah4@cornell.edu

Mark R. Rank
School of Social Work
Institute of Public Health
Campus Box 1196
Washington University
St. Louis, Missouri 63130
(314) 935-5694
markr@wustl.edu

Paper presented at the 2014 Annual Meetings of the Population Association of America, Boston, May 1 - 4.

The Social Dynamics of Economic Polarization:
Exploring the Life Course Probabilities of Top-level Income Attainment

Income polarization in the United States since 1970 is primarily the result of faster growth within the upper tail of the income distribution (Alvaredo, Atkinson, Piketty and Saez 2013; Neckerman and Torche 2007). There is no social science consensus about why income grew faster within the upper tail despite a number of investigations that have been reported over the past several decades (Grusky 2011). Much of the literature on top-level income attainment is descriptive in character, failing to indicate the social relationships associated with the process. We suspect that this attenuation in scope is related to social science's disproportionate focus on poverty and social class, to the neglect of research on the social dynamics of affluence (Massey 1996:409; Morris and Western 1999).

The present study reports analysis designed to illuminate several unexplored dimensions of top-level income concentration. First, we utilize longitudinal records within the Panel Study of Income Dynamics to identify patterns of acute versus chronic top-level income attainment. The analysis measures mobility into top-level income over the life course, correcting a possible false impression conveyed by cross-sectional research that percentile categories are synonymous over time with the occupants of those categories. The analysis identifies, for example, the percent of the total population that attains top ten percent of family income, for one or more years, between ages 25 and 60.

A second goal of the paper is to estimate whether the odds of attaining top-level income is increasing, decreasing, or remaining constant within the U.S. population. This question has relevance to the concept of social class differentiation. It is mathematically possible for income to be concentrating on an annual basis, but not in a longitudinal framework if there is sufficient mobility over time. If for example, the top ten percent of individuals is a unique set for each respective year, then there is zero net income concentration when measured over the decade. The degree that top-level income concentration is

shared longitudinally within a population effectively measures the concentration of resources, hence the degree that a society is undergoing social class differentiation.

Current Knowledge about Top-level Income Concentration

The United States has relatively high levels of wage and income inequality in comparison to other OECD countries (Kenworthy 2004; Lemieux 2008; Smeeding 2005). Income inequality within the U.S. began accelerating during the 1970s, increased markedly in the 1980s, and has continued to increase up to the present (McCall and Percheski 2010). Although there is agreement that income has concentrated at the top since 1970, measures of magnitude in the trend vary considerably. According to Current Population Survey estimates, the top 5 percent share of total household income increased from 16.6 percent in 1970 to 22.3 percent in 2010 (U.S. Census Bureau 2013, Table A-2). This finding is in contrast to a study (Saez 2013) finding that the 5 percent share increased from 20.4 in 1970 to 33.7 percent in 2010. The difference between these two findings illustrates the difference between a household survey with top-coding and a limited set income questions aimed at average income earners, versus administrative data on "taxable units" published by the U.S. Internal Revenue Service (Atkinson, Piketty and Saez 2011; Burkhauser et al. 2009). Second, that the greater quantity of income at the top reflects a qualitatively different set and social and economic relationships. Top earners are in command of different and more numerous ways making money.

There are two realms of explanation accounting for why top-level incomes have grown faster than income at other domains of the distribution. First, that relative growth in top-level income results from changing types of family income "packaging" related to marital homogamy and increased female labor force participation (McCall and Percheski 2011). Second, that relative growth in top-level income results from increased earnings inequality (Autor, Katz and Kearney 2008; Lemieux 2008; Morris and Western 1999). There is mixed evidence that income packaging is, in fact, linked to growth in family income inequality, with some studies finding empirical evidence consistent with a causal link, and others not. The

situation with regard to earnings inequality is more complex in the sense that the causal mechanism is not agreed upon. Initially it was believed that earnings inequality increases reflected "skills biased technological change" ("SBTC") where the entry of information technology into the workplace raised relative wages for higher skilled labor. Competing accounts have also been put forward including institutional change (Gabaix and Landier 2008; Mizruchi 2010) and changes in social norms (Bebchuk, Fried and Walker 2002; Kuhnen and Zwiebel 2006). It is unclear whether existing empirical evidence is capable of adjudicating between these various accounts, hence causality is unknown and is subject to continuing debate.

The present analysis is not aimed at adjudicating between different causal accounts of rising income inequality, but rather seeks to identify longitudinal dimensions of the process. At present we have found no kindred empirical studies from which to benchmark our effort.

Data and Methods

The PSID is a nationally representative, longitudinal sample of households and families interviewed annually since 1968, and constitutes the longest-running panel data set in the United States. The PSID was specifically designed to track income dynamics over time and is therefore suited for the purpose at hand. The PSID initially interviewed approximately 4,800 U.S. households in 1968 and included detailed information on roughly 18,000 individuals within those households. The PSID has since tracked these individuals annually (every other year beginning in 1997), including those children and adults who eventually broke off from their original households in order to form new households (e.g., children leaving home, new households formed as a result of separations or divorce). Thus, the PSID is designed so that in any given year the sample is representative of the entire nonimmigrant U.S. population.

The analysis tracks individuals over the prime working years from age 25 to 60. Thus individuals enter into the sample when they turn age 25, and the total number of 25-year olds over waves 1968 to

2011 is 17,043. A 25-year old entering the sample in 1968 can potentially contribute 36 person-years, whereas a 25-year old entering the sample in 2011 contributes one year. Because life table methods control for bias arising from right-censoring, both individuals contribute valid person-years to the risk set (Allison 1995).

The life table analysis employs longitudinal sampling weights in order to ensure that the PSID sample accurately reflects the U.S. population. Specifically, we utilize the longitudinal weight assigned to individuals at age 25 when they enter the life table. We utilize both the household and individual levels of information from the initial wave of 1968 through 2011. Consequently, we draw upon 44 years of longitudinal information, which translates into several hundred thousand person years of information embedded in the analysis.

Life Table Approach

Our analytical strategy is to use the household income and demographic information on individuals throughout this 44-year period in order to construct life tables that estimate the risk of top-level income across the adult lifespan. The life table examines the extent to which specific events occur across intervals of time (Namboodiri and Suchindran, 1987). In this analysis, our time intervals comprise each year (or two) that an individual ages. During that year, one can calculate the probability of an event occurring (in this case top-level income) for those who have yet to experience the event. Once the event has occurred, the individual is no longer at risk and therefore exits the life table. Based upon these age specific probabilities, the cumulative probabilities of an event occurring across the life course can be calculated. These cumulative probabilities form the core of our analysis. Individuals may contribute anywhere from 1 to 36 person-years within the life table. For example, a woman within the PSID study who turned 25 in 1975 and then in 1979 experienced a year of top-level income would have contributed five person years within our analysis. In this case, she would be included in the estimates for ages 25, 26, 27, 28, and 29. Period effects are smoothed out both within and across the age intervals.

Life tables are calculated for top 20 percent income (percentile 81 and above), top 10 percent, top 5 percent and top 1 percent. To measure chronicity for each of these levels, life tables are calculated for individuals that experience one or more years, two or more years, three or more years, four or more years, five or more years, and ten or more years. Finally, calculations are made for consecutive versus total years to identify spell length within each of these levels. For the total calculations, an individual remains in the life table risk set until he/she experiences the Nth year event for the "N or more years" category.

The percentile categories are based on weighted distributions of PSID total family income for each year of the sample, and are presented in Table 1 along with the percentile cut points from the Saez's (2013) research utilizing IRS data on taxable units. Saez does not report top 20 percent income. PSID income is sometimes top-coded, often at high levels such as $10 million. However in 1970 income was top-coded at $100,000. When top codes are encountered, we impute income by multiplying the top-code by 1.5. PSID total family income is defined as taxable income of head and wife, taxable income of others, and transfer income of head, wife and others. The PSID questionnaire includes a lengthy set of income questions designed to recover multiple forms of taxable income sources.

In general the PSID cut points are higher, reflecting more inclusive PSID total family income definition in comparison to taxable units that, depending on how taxes are filed, can split families into multiple reporting units thereby reducing the magnitude in total family income. The difference between the two series is least for the 99th percentile, perhaps because the Saez data includes a greater density of top incomes within the upper reaches of this particular category due to full coverage of administrative data vis-à-vis the PSID sample.

Growth in the magnitude of the income percentiles over time is greater within the higher up percentiles, and also greater for the PSID percentiles vis-à-vis the Saez percentiles. The foregoing sentence is based on a comparison of the first three years (1967 – 1969) versus the last three years (2006, 2008, and 2010) for the percentile cut points. The value of percentile 80 was 37 percent greater at

end of the study period compared to the beginning. For the 90th percentile, the PSID grew by 47 percent, the Saez by 36 percent; for the 95 percentile, PSID was 59 percent higher, Saez 52 percent higher; for the 99th percentile, PSID was 95 percent higher, and Saez 82 percent higher. Both of these data series reflect the same trend of income concentration within the upper reaches of the top 10 percent, and concentration with PSID family income is relatively higher compared to IRS income.

To test whether the odds of experiencing top-level income is changing over time, several sets of multivariate models are estimated where time period is entered as a covariate. The time covariate is the year that the person turns 25. This covariate represents the odds that a member of the cohort turning 25 for that year will attain top-level income. The lower hypothetical bound for this probability is equal to the value of the top-level income, the case where the same individuals attain the top-level income in every year that the cohort is in the risk set. The upper bound is the top-level category times the number of years that the cohort is at risk times the top-level category, the case representing complete year-to-year mobility. Thus for a cohort exposed to risk of attaining top 10 percent income over 10 years, the odds of attaining top level income ranges from 10 to 100 percent.

Models are estimated where time is the sole predictor variable, in addition to models that include a set of labor force, demographic, and family controls related to the causal mechanisms described in prior paragraphs. Three sets of models were estimated: Cox Proportional Hazards models, and two logit models that permit greatest flexibility with time-varying covariates. The first logit model is an event-history model where the individual exits the risk set subsequent to experiencing the event. In the second logit model sample individual remains in the risk after the event is experienced, and thus the model predicts whether the individual experienced the event for each person-year he/she is in the data. All three sets of models are estimated for weighted and non-weighted regressions, and the results are essentially all the same. In this paper we present findings from the second logit model (unweighted) because it reflects the fullest range of the data and includes the most covariates.

The covariates include education (GT 12 years of education versus LE 12 years), race (white

versus nonwhite minus Asian), age, sex, work disability (no work disability versus disability), marital status (married versus not) and presence of children (three or more children versus no children; versus 1 or 2 children). Each of these covariates is broadly consistent with a causal proposition in the literature on top-level income determination, and provides a test of whether cohort remains significant, net of a covariate vector.

Results

Table 2 reports cumulative percentages experiencing one or more years of top-level income, and demonstrates that top-level income is a common life course event. Thirteen percent of 25-year olds experience top 20 percent income, and this cumulates to 34 percent by age 30, and by age 45 two out of three individuals have experienced top 20 percent income. By age 60 nearly three quarters of the sample experienced one or more years of top 20 percent family income, suggesting high levels of income mobility experiencing this percentile level.

As anticipated the percentages experiencing respective income levels decline in relation to percentile category declines. Nevertheless, the results suggest high levels of mobility. Over half of individuals experience top 10 percent income by age 60, and 39 percent experience top 5 percent income. Twelve percent experience top one percent income between ages 25 and 60. These findings suggest that even top one percent income is relatively common when measured over the life course ages 25 to 60.

The results in Table 3 confirm the proposition that attaining top-level income is relatively prevalent in the United States. Consecutive years are less common than total years, in particular for the top one percent and top five percent levels. Attaining 10 consecutive years of top one percent is rare, and reflects the idea that only a few persist at this elite level. When compared to the 12.4 percent that ever achieve top one percent income, the findings indicate high year-to-year turnover within this category.

The multivariate results in Table 4 suggest that the odds of a sample individual experiencing top-level income is decreasing three percent per year, or thirty percent over a 10-year period. The "entry

year" coefficient is negative in for both the bivariate estimates and multivariate models. The education, labor force, race, age, and marriage and fertility are all in the expected direction with the exception of the top one percent. Having children living in with the family is a negative predictor of top twenty percent, top ten percent, and top five percent, but a positive predictor of top one percent income. This finding speaks suggests a qualitative distinction in the level of income security of the one percent versus the lower levels of income attainment.

Discussion

This study seeks to broaden social science understanding of top-level income attainment by examining mobility through these categories over time. Second, it estimates the extent that mobility into the top-level categories is increasing or decreasing over time.

A limitation of the study is the PSID over-sampling of low-income families, suggesting that the estimates of high-income families may be unreliable. However the present study utilizes all waves, generating a large sample size (N=17,043) that minimizes under sampling of high-income families.

The study has two findings. First, that top-level income attainment is far more prevalent longitudinally than in the cross-section, e.g., 12 percent of the population will be in the top one percent sometime between ages 25 and 60. Top 20 percent, top 10 percent, and top 5 percent family incomes are also more commonly experienced that their cross-sectional rates would suggest. Second, that mobility into top-level income is decreasing significantly over time, about three percent per year. Thus although top-level income in the U.S. has been relatively broadly experience, it becoming less so over the period 1968 to 2011. During this period top-level income attainment is becoming less common, and class differences are accentuating.

References

Allison, Paul. 1995. *Survival Analysis Using the SAS System: A Practical Guide*. Cary, NC: SAS Institute.

Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty and Emmanuel Saez. 2013. "The Top 1 Percent in International and Historical Perspective." *Journal of Economic Perspectives* 27:3-20,

Atkinson, Anthony B., Thomas Piketty and Emmanuel Saez. 2011. "Top Incomes in the Long Run of History." *Journal of Economic Literature* 49: 3–71.

Autor David H., Lawrence F. Katz, and Melissa S. Kearney. 2008. "Trends in U.S. Wage Inequality: Revising the Revisionists." *Review of Economics and Statistics* 90:300-323

Bebchuk, Lucian A., Jesse M. Fried and David I. Walker. 2002. "Managerial Power and Rent Extraction in the Design of Executive Compensation." *The University of Chicago Law Review* 69: 751-846

Burkhauser, Richard V., Shuaizhang Feng, Stephen P. Jenkins, and Jeff Larrimore. 2009. "Recent Trends in Top Income Shares in the USA: Reconciling Estimates from March CPS and IRS Tax Return Data." National Bureau of Economic Research Working Paper 15320.

Gabaix, Xavier and Augustin Landier. 2008. "Why has CEO Pay Increased so Much?" *The Quarterly Journal of Economics* 123:49-100.

Grusky, David B. 2011. "The Stories About Inequality That We Love to Tell." Pp. 2-15 in *The Inequality Reader: Contemporary and Foundational Readings in Race, Class and Gender* edited by David B. Grusky and Szonja Szelenyi. Boulder, CO: Westview Press.

Kenworthy, Lane. 2004. "Inequality and Sociology." American Behavioral Scientist 50:584-602.

Kuhnen, Camelia M. and Jeffrey Zwiebel. 2006. "Executive Pay, Hidden Compensation, and Managerial Entrenchment." Working Paper, Northwestern University and Stanford University.

Lemieux, Thomas. 2008. "The Changing Nature of Wage Inequality." *Journal of Population Economics* 21:21-48.

Massey, Douglass S. 1996. "The Age of Extremes: Concentrated Affluence and Poverty in the Twenty-First Century." *Demography* 33:395-412.

McCall, Leslie and Christine Percheski. 2010. "Income Inequality: New Trends and Research Directions." *Annual Review of Sociology* 36:329-347.

Mizruchi, Mark S. 2010. "The American Corporate Elite and the Historical Roots of the Financial Crisis of 2008." *Markets on Trial: The Economic Sociology of the U.S. Financial Crisis: Part B. Research in the Sociology of Organizations* 30B:103-139.

Morris, Martina and Bruce Western. 1999. "U.S. Earnings Inequality at the Close of the Twentieth Century." *Annual Review of Sociology* 25:623-657.

Namboodiri, Krishnan and C. M. Suchindran. 1987. *Life Tables and Their Applications*. Orlando, FL: Academic.

Saez, Emmanuel. 2013. "Striking it Richer: The Evolution of Top Incomes in the United States (Updated with 2012 Preliminary Estimates)." Unpublished paper, Retrieved March 8, 2014.http://elsa.berkeley.edu/users/saez/#income.

Smeeding, Timothy M. 2005. "Public Policy, Economic Inequality, and Poverty: The United States in Comparative Perspective." *Social Science Quarterly* 86(5): 955–983.

U.S. Census Bureau. 2013. Current Population Reports P60-245, Income, Poverty, and Health Insurance Coverage in the United States: 2012. Washington, DC: U.S. Government Printing Office.

Table 1: Affluent percentile thresholds, PSID versus Saez[1,2]

	Percentile 80	Percentile 90		Percentile 95		Percentile 99	
Year	PSID	PSID	Saez	PSID	Saez	PSID	Saez
1967	72,881	94,168	77,123	113,360	95,550	171,737	187,463
1968	77,087	100,589	80,245	122,211	98,750	191,151	193,396
1969	82,010	103,583	82,565	128,067	102,141	193,437	191,666
1970	80,944	103,817	82,894	125,913	103,131	191,119	189,273
1971	83,120	107,704	83,425	131,937	103,926	200,383	191,138
1972	87,136	109,572	87,225	131,695	107,487	217,578	198,602
1973	88,419	113,225	88,833	135,566	110,829	226,942	205,047
1974	85,656	110,598	86,777	136,257	110,492	222,524	201,495
1975	82,801	105,401	84,032	132,360	104,894	213,235	189,564
1976	87,968	112,837	85,949	139,116	106,977	234,114	191,433
1977	89,497	115,852	86,888	140,361	108,366	217,740	192,344
1978	91,244	117,078	88,333	143,839	109,859	229,473	196,533
1979	90,124	117,161	88,080	141,329	109,216	228,014	194,667
1980	86,817	111,962	87,178	138,644	107,936	222,287	192,284
1981	84,217	108,450	86,263	135,647	107,193	216,483	185,508
1982	82,042	108,711	85,786	133,346	105,652	211,998	185,628
1983	86,592	113,867	85,398	139,378	105,554	219,320	183,584
1984	89,843	115,452	87,765	147,107	109,413	235,775	191,349
1985	90,453	119,590	89,213	147,866	111,601	229,477	197,561
1986	94,085	122,582	91,045	152,845	113,339	240,388	200,308
1987	95,803	123,641	93,481	155,127	116,934	254,770	221,854
1988	97,712	128,500	95,066	162,976	121,061	277,464	244,599
1989	91,055	119,413	95,557	151,263	123,289	257,154	247,643
1990	90,864	118,532	94,655	148,598	122,169	257,815	247,278
1991	88,238	116,896	94,487	145,240	123,132	248,205	241,726
1992	93,272	123,585	94,474	157,318	122,587	262,715	251,517
1993	98,482	130,401	94,224	163,900	121,967	333,888	251,693
1994	98,750	133,188	95,720	169,245	125,002	308,292	258,468
1995	98,111	133,088	98,054	173,441	128,498	301,423	268,272
1996	99,111	134,912	97,235	174,313	131,616	320,269	283,745
1998	107,712	147,184	103,689	194,015	141,122	364,160	313,245
2000	112,723	154,156	107,547	209,234	147,922	440,253	331,811
2002	109,353	145,966	105,779	192,197	143,361	363,702	308,631
2004	109,685	149,823	105,424	195,124	144,090	362,315	326,014
2006	108,998	150,214	109,300	202,089	151,582	379,511	356,216
2008	106,611	147,696	108,983	193,332	151,202	376,608	351,788
2010	102,090	140,800	107,023	181,000	147,933	332,300	335,915

1. Values reflect 2010 prices, computed using CPI-U.
2. The PSID percentiles are computed by the authors. The Saez percentiles are from Saez (2013).

Table 2. Cumulative Percentage of American Adults Experiencing Various Levels of Household Affluence by Age

	Top of Income Distribution			
Age	20 percent	10 percent	5 percent	1 percent
25	13.1%	6.3%	3.7%	0.7%
30	33.9%	17.7%	9.5%	1.9%
35	48.9%	27.6%	15.9%	3.6%
40	58.7%	36.8%	22.6%	5.9%
45	66.5%	44.5%	28.6%	7.9%
50	70.3%	50.9%	33.1%	9.6%
55	72.6%	54.8%	37.4%	11.4%
60	73.4%	56.4%	39.3%	12.4%

Table 3. Years of Affluence Experienced Between the Ages of 25 to 60

	Top of Income Distribution			
Years of Affluence	20 percent	10 percent	5 percent	1 percent
Total Years				
1 or more	73.4%	56.4%	39.3%	12.4%
2 or more	64.5%	44.5%	28.6%	7.4%
3 or more	60.6%	39.1%	23.8%	4.4%
4 or more	55.2%	34.6%	19.5%	3.5%
5 or more	52.5%	29.4%	15.3%	2.9%
10 or more	37.0%	17.5%	8.1%	1.1%
Consecutive Years				
1 or more	73.4%	56.4%	39.3%	12.4%
2 or more	60.5%	39.4%	24.8%	5.3%
3 or more	53.0%	32.0%	17.7%	3.3%
4 or more	46.5%	27.1%	13.2%	2.5%
5 or more	40.8%	20.8%	9.1%	1.6%
10 or more	23.0%	9.7%	4.5%	0.6%

Table 4: Logistic regression analysis of affluence: odds ratios[1]

	Top 20 percent		Top 10 percent	
Covariates	I	II	I	II
Entry year	0.967	0.971	0.966	0.972
Education GT 12 (LE 12)		2.473		2.468
Race = white		2.14		2.404
Age		1.04		1.042
Sex = male		1.014[NS]		1.017[NS]
No work disability		2.753		2.449
Married (not married)		6.396		5.738
(GE 3 children)				
0 children		1.55		1.497
1 or 2 children		1.293		1.269
Intercept[2]	64.9	51.1	67.0	48.2
-2LL	187,773	149,067	118,185	96,339
df	1	9	1	9

	Top 5 percent		Top 1 percent	
Entry year	0.965	0.973	0.965	0.975
Education GT 12 (LE 12)		2.498		3.006
Race = white		3.166		7.444
Age		1.041		1.054
Sex = male		1.00[NS]		0.882*
No work disability		2.329		2.125
Married (not married)		5.028		3.511
(GE 3 children)				
0 children		1.402		0.583
1 or 2 children		1.161		0.754
Intercept	67.7	45.7	66.1	39.4
-2LL	69,899	58,110	17,802	15,070
df	1	9	1	9

1. All odds ratios are statistically significant at $p < .001$, unless otherwise indicated.
2. The intercept is a coefficient, not an odds ratio.
*p<.05