# How Risky Are Recessions for Top Earners?

*By* Fatih Guvenen, Greg Kaplan and Jae Song[*]

This paper addresses two open questions about the fortunes of top earners:

1) How sensitive are the earnings of top earners to business cycle fluctuations?

2) How does the business cycle sensitivity of top earners vary by industry?

To answer these questions, we use a confidential panel data set on earnings histories from the U.S. Social Security Administration (SSA), which allows us to follow workers' earnings over the business cycle. Our approach is to identify individuals who were in the top 1 percent of the earnings distribution just prior to the economy entering a recession or boom. We ask how the evolution of these workers' subsequent earnings compares with the evolution of workers' earnings from the rest of the population, and how these differences vary by sector.[1]

Despite the extensive recent attention given to top earners, surprisingly little is known about the answers to these two questions, largely due to the lack of detailed panel data with sufficient coverage at the top of the earnings distribution. Panel data is necessary to study the dynamics inherent in these questions. However, most survey based datasets under-represent very high earners, because of the extremely low response rate of such individuals to surveys, and because of top-coding in many frequently used panel datasets. The contribution of this paper is to use a very large panel dataset from administrative sources with non-top-coded earnings information that allows us to track top earners from 1978 to 2011, covering three full expansions and four full recessions.

In the absence of panel data on top earners, earlier papers have characterized the evolution of top income shares. This is accomplished by computing the share of aggregate income earned by the top 1 percent of individuals year by year, and documenting the changes in this quantity over time. Of course, the group of top earners is not comprised of the same individuals in every year. Hence compositional changes can confound inference.

Parker and Vissing-Jorgensen (2010) and Saez (2013) are two prominent examples of this approach. Both of these papers conclude that the income shares of top earnings groups have been highly cyclical since the 1980s. While these are interesting findings, they do not directly answer the questions raised above, because of the lack of a panel dimension to the analyses, and the lack of disaggregated data by sector.[2] Bakija, Cole and Heim (2012) study the occupations of top earners using tax data on earnings and self-reported occupations. They have annual data on occupation from 2001 to 2005, but only have four data points between 1979 to 2000. Our data contain annual information on industry and earnings for all years between 1978 and 2011. Finally, Guvenen, Ozkan and Song (2013) have documented that individuals who enter a recession in the top 1 and 0.1 percent of the income distribution face the largest earnings drops of any group in the population. But they have not analyzed the data

[*] Guvenen: University of Minnesota, Minneapolis, MN 55455, and NBER (email: guvenen@umn.edu). Kaplan: Princeton University, Princeton, NJ 08544 and NBER (email: gkaplan@princeton.edu). Song: Social Security Administration (email: jae.song@ssa.gov). The views expressed in this paper are those of the authors and not necessarily those of the Social Security Administration.

[1] In this paper, we look only at the top 1 percent. Guvenen, Kaplan and Song (2014) contains a more detailed analysis that also covers the top 0.1 percent and 0.01 percent.

[2] Parker and Vissing-Jorgensen (2010) provide some decompositions by broad sectors, but only using data from the Current Population Survey, which is top-coded and under-samples the very rich. They also undertake an analysis using panel data, but only for Canada.

by sectors as is done here.

# I.  Data

We use confidential data on individual earnings histories from SSA records. Our basic source of information is the Master Earnings File (MEF), a panel dataset covering all US individuals with a Social Security Number. We work with a 10 percent representative sample of U.S. males from the MEF, covering the period 1978 to 2011. Earnings data come directly from individual W-2 forms (Box 1) and, as such, include wages and salaries, bonuses, and exercised stock options. In addition, the dataset contains demographic characteristics, such as date of birth, sex, and race.

Earnings data in the MEF are not capped, which is crucial for our analysis of top earners, but also means that the MEF contains a small number of extremely high earnings observations each year ($>$ $\$100,000,000$). To avoid such outliers unduly influencing average facts about the top 1 percent, we cap (winsorize) observations above the 99.999th percentile. We deflate nominal earnings using the Personal Consumption Expenditure deflator, and express all quantities in 2005 dollars.[3]

Our analysis is distinguished from previous analyses of top earners by the panel dimension of the data. We exploit this panel dimension in two ways. First, we define membership of the top 1 percent based on average earnings over a 5-year window, rather than earnings in a single year. We thus focus on those individuals who are persistent top earners, and exclude individuals whose high earnings are a one-off event. This distinction is important given the relatively high probabilities of transiting in and out of the annual top 1 percent documented in Kopczuk, Saez and Song (2010), Auten, Gee and Turner (2013) and Guvenen, Kaplan and Song (2014). Second, and more importantly, we compute statistics on the earnings growth of individuals in the top 1 percent, rather than the growth in average earnings of the top 1 percent. This distinction is important because of compositional changes that arise due to the aforementioned instability of membership of this group.

Specifically, we construct a rolling panel, whereby for a given year $t$, we select a group of individuals that satisfy certain selection criteria during years $t-5$ to $t-1$. We then compute the earnings growth experienced by these individuals between years $t$ and $t+5$. An individual is eligible for inclusion in the rolling panel for year $t$ if he satisfies the following two conditions in three of the previous five years, including year $t-1$: (i) he must be between the ages of 25 and 60; and (ii) his earnings must satisfy a minimum threshold.[4] We compute the average earnings from $t-5$ to $t-1$ for all such eligible workers, and assign each individual to either the top 1 percent or bottom 99 percent based on their position in this distribution of average earnings. Let $D_{i,t}^{1pc} = 1$ if individual $i$ is in the top 1 percent in year $t$ according to this definition. The $t-5$ to $t-1$ average earnings thresholds for membership of the top 1 percent ranged from a minimum of around $\$217,000$ in $t = 1984$ to a maximum of around $\$348,000$ in $t = 2005$.

Our dataset also contains information on the sector in which an individual works. These data come from linking Employer Identification Numbers, which are available on the W-2 form, to 5-digit standard industry classification (SIC) codes. We aggregate these to standard 1-digit industry classifications. Where an individual has multiple employers, we use the one that corresponds to the highest income W-2. For each rolling panel at year $t$, we assign individuals to sectors based on their SIC code in year $t-1$, which is the most recent year of the 5-year period used to construct average earnings.

---

[3]Our construction of the 10 percent sample follows Guvenen, Ozkan and Song (2013). That paper also contains further information about the MEF and justification for the sample selection choices made in this paper.

[4]The minimum threshold is equal to one-half of the legal minimum wage times 520 hours (13 weeks at 40 hours per week). A condition of this kind is standard in the literature on income dynamics. However, because the main focus of our analysis are top earners, varying this minimum threshold is unlikely to affect our results.
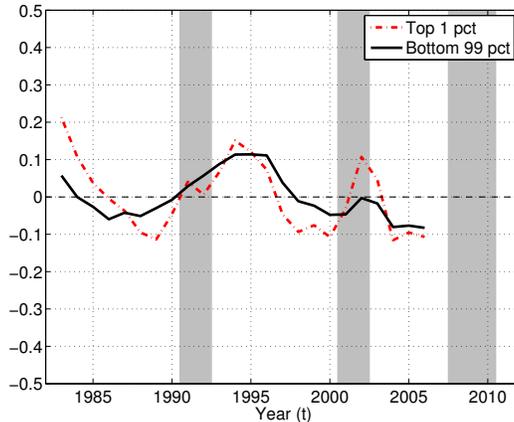
FIGURE 1. CYCLICALITY OF 5-YEAR EARNINGS GROWTH

Notes: Plot shows average growth in log earnings from $t$ to $t+5$, conditional on rank in distribution of average earnings from $t-5$ to $t-1$. Lines are normalized to have mean zero over the sample period.

## II.  Cyclicality of top earners

Organizing the data in this way allows us to follow the subsequent fortunes of individuals, conditional on their recent history, across recessions and expansions. Figure I illustrates these fluctuations by plotting the average 5-year forward change in log earnings, conditional on membership of either the top 1 percent or bottom 99 percent based on average earnings over the previous 5 years. That is, the figure plots $E\left[\Delta_5 \log y_{i,t+5}|D_{i,t}^{1pc}\right]$.[5] To understand the interpretation of this graph, consider, for example, the 1989 data point on the red dashed line. This shows that the average income change between 1989 and 1994, for individuals who were in the top 1 percent of the distribution of average earnings from 1983 to 1988, was a drop of just over 10 log points.

Figure I reveals substantial business cycle fluctuations in average 5-year earnings growth for both the top 1 percent and bottom 99 percent. However, in contrast with the findings using cross-sectional data in Parker and Vissing-Jorgensen (2010) and Saez (2013), the panel data evidence, as illustrated in Figure I, suggests that the

two groups experience similarly sized cyclical fluctuations in earnings growth. More concretely, the time-series standard deviation of average earnings growth is somewhat larger for the top 1 percent than the bottom 99 percent (0.095 versus 0.062), but as we shall see, this additional volatility is unrelated to the business cycle.

To measure the cyclicality of earnings more formally, we regress individual log earnings growth from $t$ to $t+5$ on a constant, a linear time trend and the growth in GDP between $t$ and $t+5$.

$$(1) \qquad E\left[\Delta_5 \log y_{i,t+5}|D_{i,t}^{1pc}=k\right]$$
$$= \alpha^k + \gamma^k t + \beta^k \Delta_5 GDP_{t+5} + \varepsilon_t^k.$$

We estimate equation (1) separately for the top 1 percent $(k=1)$ and the bottom 99 percent $(k=0)$. The coefficient $\beta$ measures the cyclicality of average earnings growth, the constant $\alpha$ is included to capture potential mean reversion (which is expected, since we are conditioning on a group of individuals with very high earnings in the previous five years), and the coefficient $\gamma$ captures a potential time trend in earnings growth.

For the period $t = 1983 \ldots 2006$ we estimate a loading factor for the top 1 percent of $\hat{\beta} = 1.55$, with a standard error of 0.32. For the bottom 99 percent the cor-

---

[5]When computing statistics for the bottom 99 percent, we use a 2 percent random subsample of this group, since the full sample is very large, consisting of around 5 million individuals per year.

Table 1—Summary statistics by sector

| | Top 1 Percent | | | Bottom 99 Percent | | |
|---|---|---|---|---|---|---|
| | Percent | 5 yr av. earns | St Dev $\Delta_5 y$ | Percent | 5 yr av. earns | St Dev $\Delta_5 y$ |
| Mining | 1.7 | 534,492 | 0.172 | 2.7 | 44,185 | 0.072 |
| Construction | 3.4 | 503,243 | 0.209 | 8.0 | 42,890 | 0.118 |
| Manufacturing | 17.4 | 559,877 | 0.131 | 22.2 | 55,684 | 0.067 |
| Transportation, Utilities | 3.4 | 522,660 | 0.166 | 8.6 | 56,606 | 0.063 |
| Wholesale Trade | 7.4 | 549,685 | 0.101 | 5.9 | 51,637 | 0.071 |
| Retail Trade | 4.2 | 570,874 | 0.120 | 9.3 | 38,559 | 0.083 |
| FIRE | 18.2 | 631,445 | 0.166 | 5.3 | 63,687 | 0.081 |
| Services | 39.4 | 549,469 | 0.099 | 23.3 | 51,154 | 0.056 |
| Nonclassifiable | 3.0 | 567,317 | 0.128 | 3.1 | 46,067 | 0.083 |
| All Sectors | | 566,607 | 0.095 | | 50,850 | 0.062 |

*Notes:* Statistics are averages from 1983-2006. Statistics for individuals with missing SICs are not reported. Statistics for individuals with SIC codes 100-999, 9100-9899 and 4300-4399 (agriculture and government) are not reported due to small number of individuals in top 1%. Hence reported percentages do not add to 100%.

responding estimate is $\hat{\beta} = 1.31$, with a standard error of 0.14. Thus for the two groups as a whole, we find a very similar sensitivity of earnings to aggregate fluctuations. Moreover, the explained variation from this regression is $R^2 = 0.60$ for the top 1 percent and $R^2 = 0.82$ for the bottom 99 percent. Hence the additional earnings volatility of top earners is largely uncorrelated with the business cycle. In summary, on average individuals in the top one percent do not experience much larger cyclical earnings fluctuations than the rest of the population. The next section shows that the same is not true once we condition on particular industries.

### III. Cyclicality of top earners by sector

Summary statistics by sector are reported in Table II. First, observe that the largest group in the top 1 percent is Services (39.4%), consistent with its representation in the overall economy (23.3%). Second, observe that the most over-represented group among top earners is FIRE (Finance, Insurance, Real Estate) which makes up 18.2% of the top 1 percent and only 5.3% of the bottom 99 percent. Third, there is substantial variation in average earnings across industries even within the top 1 percent: workers in FIRE have the highest average annual earnings (around $630,000) while those in Construction have the lowest (around $500,000). Table II also reports the time-series standard deviation of average 5-

year earnings growth by sector. There are substantial differences in earnings growth volatility across sectors: Construction is the most volatile sector while Services is the least volatile sector.

A large part of these differences in volatility across sectors pertains to business cycle fluctuations. We illustrate this feature of the data graphically in Figure III by reproducing Figure I separately for top earners in each sector $j$. That is, the lines in Figure III plot $E\left[\Delta_5 \log y_{i,t+5} | D_{i,t}^{1pc} = 1, \text{SIC} = j\right]$. The left panel, Figure 2(a), shows average earnings growth for the four sectors with the highest volatility, while the right panel, Figure 2(b), shows the three sectors with the least volatility. The figure illustrates clearly that top earners in Construction, Manufacturing, Transportation and Utilities, and FIRE experience substantial changes of fortunes from recessions to expansions, while top earners in Services and Wholesale Trade show only modest fluctuations. Comparing Figure I and Figure 2(b) reveals that top earners in Services (nearly 40 percent of all top earners) exhibit fluctuations that are roughly the same size as the average earner in the bottom 99 percent.

To explore the differences in business cycle fluctuations across sectors more formally, we re-estimate the regression in (1) separately by sector and top earner status. Allowing for differential time trends across sectors is important because of the secular trends in earnings growth that have

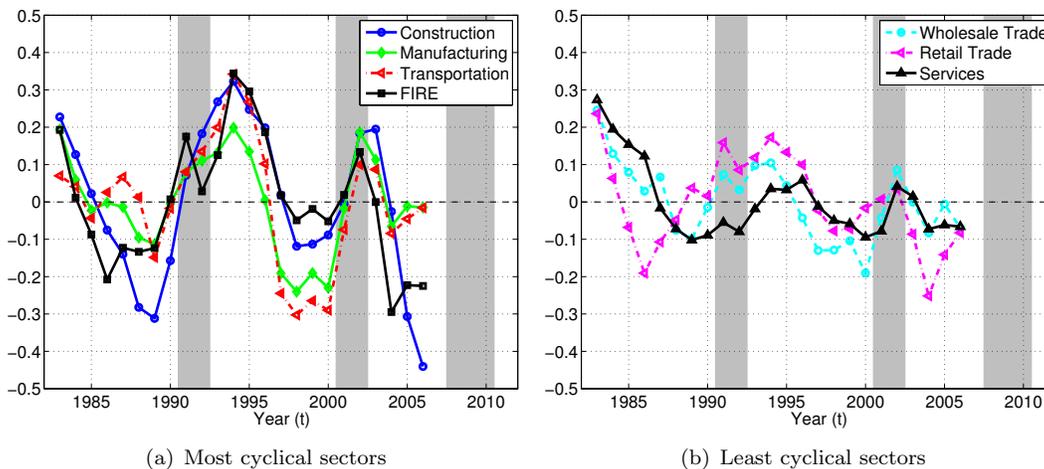(a) Most cyclical sectors

(b) Least cyclical sectors

FIGURE 2. MEAN 5-YEAR EARNINGS GROWTH OF INDIVIDUALS IN TOP 1 PERCENT, BY SECTOR

been observed during this time period., e.g. top earners in FIRE have seen accelerating earnings growth relative to the average, while Manufacturing has seen decelerating growth.

These findings, which form the key results of the paper, are summarized in Table III. The first three columns report the estimates $\hat{\beta}$ together with standard errors and the associated $R^2$ from regression (1), for the top 1 percent of earners. The second three columns report the analogous statistics for the bottom 99 percent. The key findings are neatly illustrated by considering the two largest sectors among the top 1 percent: Services and FIRE. Services, which contains the biggest proportion of top earners, displays the least amount of systematic risk for top earners, with an estimated loading factor of $\hat{\beta} = 0.61$ and an explained variation of only $R^2 = 0.36$. Top earners in FIRE, on the hand, experience extremely cyclical earnings growth: the loading factor is $\hat{\beta} = 3.63$ and the explained variation is $R^2 = 0.85$. Moreover, aside from FIRE and Construction, the remaining sectors have very low explained variation for workers in the top 1 percent.

For workers in the bottom 99 percent, the biggest sector, Services, is also the least cyclical, with an estimated loading factor of $\hat{\beta} = 1.03$ and an explained variation of only

$R^2 = 0.72$ (but note that Service workers in the top 1 percent have less cyclical earnings growth than service workers in the bottom 99 percent). Like the top 1 percent, Construction workers have the most cyclical earnings growth among workers in the bottom 99 percent ($\hat{\beta} = 2.58$, $R^2 = 0.90$). But unlike top earners, workers in the bottom 99% in FIRE do not have particularly cyclical earnings growth.

## IV. Conclusions

Our findings suggest large differences across sectors in the cyclicality of earnings growth. While average earnings growth (across workers in all sectors) is very similar for individuals in the top 1 percent compared with the bottom 99 percent, there are striking differences across sectors. Most of the cyclicality of top earners is driven by those in FIRE, while FIRE workers in the rest of the distribution do not have particularly cyclical earnings. Services on the hand, the largest sector, is actually less cyclical in the top 1 percent than in the bulk of the earnings distribution.

The analysis in this paper has only scratched the surface of issues surrounding the nature of the business cycle risk faced by top earners. Guvenen, Kaplan and Song (2014) use an even larger sam-

TABLE 2—CYCLICALITY OF LABOR EARNINGS BY SECTOR

| | Top 1 Percent | | | Bottom 99 Percent | | |
|---|---|---|---|---|---|---|
| | $\hat{\beta}^j$ | s.e.$(\hat{\beta}^j)$ | $R^2$ | $\hat{\beta}^j$ | s.e.$(\hat{\beta}^j)$ | $R^2$ |
| Mining | 1.91 | 0.59 | 0.58 | 1.34 | 0.20 | 0.72 |
| Construction | 4.15 | 0.60 | 0.70 | 2.58 | 0.20 | 0.90 |
| Manufacturing | 1.15 | 0.63 | 0.17 | 1.24 | 0.19 | 0.70 |
| Transportation, Utilities | 1.41 | 0.80 | 0.18 | 1.09 | 0.23 | 0.53 |
| Wholesale Trade | 0.65 | 0.43 | 0.35 | 1.43 | 0.17 | 0.79 |
| Retail Trade | 2.11 | 0.37 | 0.66 | 1.69 | 0.21 | 0.78 |
| FIRE | 3.63 | 0.34 | 0.85 | 1.69 | 0.19 | 0.81 |
| Services | 0.61 | 0.42 | 0.36 | 1.03 | 0.16 | 0.72 |
| Nonclassifiable | 1.23 | 0.60 | 0.21 | 1.75 | 0.22 | 0.76 |
| All Sectors | 1.55 | 0.32 | 0.60 | 1.31 | 0.14 | 0.82 |

*Notes:* Estimates of equation (1) for $t = 1983\ldots2006$. Results for individuals with missing SICs are not reported. Results for individuals with SIC codes 100-999, 9100-9899 and 4300-4399 (agriculture and government) are not reported due to small number of individuals in top 1%.

ple from the same panel dataset that allows them to study individuals in the top 0.1% and 0.01%, as well as to distinguish between several two-digit SIC industries: for example, separating health services (mostly doctors), and professional services (lawyers, engineers, accountants, various research services, etc) from the broad category of services, as well as distinguishing between finance, insurance and real estate. Such distinction is important, especially because these categories occupy much larger shares of the top 1% and 0.1% than their population share. This larger dataset also enables further exploitation of the panel dimension to study membership in the top percentiles of average earnings over horizons that are longer than 5 years, such as lifetime earnings.

**REFERENCES**

**Auten, Gerald, Geoffrey Gee, and Nicholas Turner.** 2013. "Income Inequality, Mobility, and Turnover at the Top in the US, 1987-2010." *American Economic Review*, 103(3): 168–72.

**Bakija, Jon, Adam Cole, and Bradley T Heim.** 2012. "Jobs and Income Growth of Top Earners and the Causes of Changing Income Inequality: Evidence from US Tax Return Data." *Mimeo, Williams College.*

**Guvenen, Fatih, Greg Kaplan, and Jae Song.** 2014. "The Life Cycle and Business Cycle of Top Earners." *Mimeo, University of Minnesota.*

**Guvenen, Fatih, Serdar Ozkan, and Jae Song.** 2013. "The Nature of Countercyclical Income Risk." *Journal of Political Economy.*

**Kopczuk, Wojciech, Emmanuel Saez, and Jae Song.** 2010. "Earnings inequality and mobility in the United States: evidence from social security data since 1937." *The Quarterly Journal of Economics*, 125(1): 91–128.

**Parker, Jonathan A, and Annette Vissing-Jorgensen.** 2010. "The Increase in Income Cyclicality of High-Income Households and Its Relation to the Rise in Top Income Shares [with Comments and Discussion]." *Brookings Papers on Economic Activity*, 1–70.

**Saez, Emmanuel.** 2013. "Striking it Richer: The Evolution of Top Incomes in the United States." *Mimeo, University of California, Berkeley.*