Wages and Human Capital in the U.S. Finance Industry: 1909–2006

ONLINE APPENDIX

Thomas Philippon* and Ariell Reshef†

May 2012

A Historical wage data

The data are from the Bureau of Economic Analysis Industry Accounts, Martin (1939) and Kuznets (1941). Kuznets (1941) gives estimates of net income, wages and salaries, and the number of employees separately for banking, insurance, and real estate over the period 1919–1938. The banking category, however, covers only commercial banks, savings banks, and federal reserve banks; brokerage, investment banking, and other financial activities are not included. As a result, the size of the industry is smaller than that implied by BEA data. Fortunately, there is large overlap of 10 years with the BEA data over which the correlation between the two series is 96.6%. It seems therefore quite safe to impute values for the period 1919–1928 using Kuznets' data.

Martin (1939) provides data for finance, insurance, and real estate, but not for finance and insurance only. For the period 1909–1929 the estimates are based on data collected from banking, insurance, and real estate. For the period 1899–1908, however, the 1909 estimate is "projected to 1899 on the basis of other data indicating a probable trend for this period." We find this procedure questionable, and therefore truncate our sample in 1909. We collected additional data from Mitchell (1921) for the banking sector in 1909–1919. The implied banking wage calculated from Mitchell (1921) is quite similar to the implied wage from Martin (1939) and the census data to measure the number of employees, except that it grows slightly faster.

As mentioned, the data from Martin (1939) includes real estate. This does not appear to be a problem for the long-run trends. Using BEA data for the period 1929–2005, we find a correlation of 0.993 between the relative wage series including real estate and that excluding real estate.

B BEA data sources

B.1 Industry Accounts

The industry accounts are prepared by the BEA's Current Industry Analysis Division. We use the NAICS classification for "Compensation of employees" (wages and salaries, and supplements) and for "Full-time equivalent employees" in 1987–2006. We use the SIC classification in 1947–1987, which itself changes in 1972. We use Tables 6.2A and 6.5A from the BEA's Income and Employment by Industry tables in 1929–1946. Mapping the data before

*Stern School of Business, New York University; NBER and CEPR.
†University of Virginia.

and after 1946 requires adjusting for changes in the classification of real estate activities. We report here the relevant information on methodology from Bureau of Economic Analysis (2011a).

Compensation of employees

"Compensation of employees, paid" (NW) is the income accruing to employees as remuneration for their work for domestic production; it includes compensation paid to the rest of the world and excludes compensation received from the rest of the world. It is the sum of wage and salary accruals and of supplements to wages and salaries.

"Wage and salary accruals" consist of the monetary remuneration of employees, including the compensation of corporate officers; commissions, tips, and bonuses; voluntary employee contributions to certain deferred compensation plans, such as 401(k) plans; employee gains from exercising nonqualified stock options; receipts-in-kind; and miscellaneous compensation of employees. Miscellaneous compensation of employees includes judicial fees paid to jurors and to witnesses, compensation of prison inmates, and marriage fees paid to justices of the peace. Wage and salary accruals consist of disbursements and wage accruals less disbursements. Disbursements are wages and salaries as just defined except that retroactive wage payments are recorded when paid rather than when earned. Accruals less disbursements is the difference between wages earned, or accrued, and wages paid, or disbursed. In the NIPAs, wages accruals is the measure used for gross domestic income, and wage disbursements is the measure used for personal income.

"Supplements to wages and salaries" consist of employer contributions for employee pension and insurance funds and of employer contributions for government social insurance.

Full-time equivalent employees

Full-time equivalent employees (FTE) equal the number of employees on full-time schedules plus the number of employees on part-time schedules converted to a full-time basis. The number of full-time equivalent employees in each industry is the product of the total number of employees and the ratio of average weekly hours per employee for all employees to average weekly hours per employee on full-time schedules. An industry's full-time equivalent employment will be less than the number of its employees on full- and part-time schedules, unless it has no part-time employees.

The FTE in sector s and for time t is given by

$$FTE_{s,t} = N_{s,t} \cdot \frac{\overline{hrs}_{s,t}}{\overline{hrs}_{ft,s,t}} = N_{s,t} \cdot \frac{1}{\theta_{s,t}} \ ,$$

where N is employment (bodies), $\overline{hrs}$ is the average number of weekly hours of all workers, $\overline{hrs}_{ft}$ is the average number of weekly hours of full-time workers, and $\theta_{s,t} = \overline{hrs}_{ft,s,t}/\overline{hrs}_{s,t} = N_{s,t}/FTE_{s,t}$ is the full-time equivalence correction. As long as there are some part-time employees, $\theta_{s,t} > 1$. Since the BEA reports both $N_{s,t}$ and $FTE_{s,t}$, we obtain $\theta_{s,t}$ by taking their ratio.

B.2 State Personal Income

The Annual State Personal Income (SPI) tables are published by the BEA's Regional Income Division. We report here the relevant information on methodology from Bureau of Economic Analysis (2011b).

For employment we use the "full and part time employment by industry" series. For worker compensation we use the "earnings by industry" series. While the earnings data by state and industry are available from 1958 on, the employment data are available only from 1969 on. The data are organized along the NAICS from 1990 on, with an overlap with the SIC in 1969–2001. This allows for adjustments in the series for changes in the classifications. The components of earnings are "wage and salary disbursements", "supplements to wages and salaries", and "proprietors' income".

"Wage and salary disbursements" consist of the monetary remuneration of employees, including corporate officers' salaries and bonuses, commissions, pay-in-kind, incentive payments, and tips. They reflect the amount of payments disbursed but not necessarily earned during the year. Wage and salary disbursements are measured before deductions, such as social security contributions and union dues. Wage and salary disbursements include stock options of nonqualified plans at the time they are exercised. Stock options are reported in wage and salary disbursements. The value that is included in wages is the difference between the exercise price and the price at which the stock options were granted.

Supplements to wages and salaries include employer contributions for employee pension and insurance funds and for government social insurance. Proprietors' income excludes dividends and monetary interest received by non-financial business and rental incomes received by persons not primarily engaged in the real estate business; these incomes are included in dividends, net interest, and the rental income of persons, respectively.

Personal income by state is based on a "place of residency" concept. The BEA adjusts earnings by place of work for commuters' income to reflect the place of residency. However, these data are not provided by industry. The data we use are based on a "place of work" concept, that is, where the income is generated, which is not necessarily where the individual resides.

C Current Population Survey

C.1 March supplement cross section

Our data on individuals come from the March supplement of the Current Population Survey (CPS) (Annual Social and Economic Study) from survey years 1968–2006, which pertain to the actual years 1967–2005. A CPS year refers to data of the preceding year; that is March CPS 2006 documents annual data from calendar year 2005. We therefore adopt the following taxonomy: We call the actual year that the survey pertains to as, simply, "year", while a CPS year is denoted as the survey year. The CPS is a monthly survey of about 50,000 households conducted by the U.S. Census Bureau for the U.S. Bureau of Labor Statistics. Currently, there are more than 65,000 participating households. The sample is selected to represent the civilian non-institutional U.S. population. The CPS includes data on employment, unemployment, earnings, hours of work, and other demographic characteristics, including age, sex, race, marital status, and educational attainment. Also available are data on occupation, industry, and class of worker. We choose to use only one particular month's survey, the March supplement, for two reasons. First, this supplement contains more demographic details, particularly on work experience and income sources and amounts. Since 1976, the survey has also been supplemented with a sample of Hispanic households (about 2,500 interviewed). Second, it has been extensively used in the empirical labor and macro-labor literature, which lends to the comparability of our results. We now define the groups used in our empirical analysis. We restrict attention to individuals in the labor force, who are at least 15 years of age.

Occupations
Examining the distribution of occupations within finance and its three subsectors leads us to choose seven occupational groups (henceforth occupations) that describe the major occupational groups in our sample: managers and professionals; mathematics and computers; insurance specialists (insurance sales persons, statisticians and actuaries); brokers and traders; bank tellers; administration (including clerks); and all the rest (janitors, security and miscellaneous). As with industry classifications, major occupational reclassifications occur in survey year 1983, from the census 1970 system to the 1980 system, and in survey year 2003, from the census 1990 system to the 2000 system. Of these two reclassifications, the latter was more substantial. We examined the occupational crosswalks, provided by the Census Bureau, to make sure that our occupational groups are consistently defined over time (Census Bureau (1989), Census Bureau (2003)). Our criteria for grouping occupations under one title are stability in occupational shares and in relative wages. In some cases we could not consistently separate managers from professionals due to reclassifications in survey years 1983 and 2003; some occupations that were defined as professional were split and reclassified as managerial and vice versa. However, these two groups together are consistently identified, without any jumps or drops in their employment shares over time or in their relative wages. Much effort was devoted to ensuring that the other occupation groups are also consistently defined throughout our sample. Note that some of these occupations potentially mean different things in different industries. For instance, in credit intermediation managers and professionals include bank officers, but these officers do not exist in the two other industries. The composition of administration occupations also varies across subsectors of finance. However, our more narrowly defined occupations—mathematics and computers, insurance specialists, brokers and traders, and bank tellers—are consistently defined.

Industry classification
The financial sector includes three industries: credit intermediation, insurance, and other finance. To define the private sector, we exclude all government employees, as well as employees of the U.S. Postal Service. Banks

and thrift and saving institutions are included in credit intermediation. Securities, commodities, funds, trusts, and other financial investments, as well as investment banks are all included in other finance. These sectors are consistently identified, without any jumps or drops in their shares of total employment, despite changes in industrial classifications in the CPS in our sample, which occur following each decennial census. Major industrial reclassifications occur in survey year 1983, from the census 1970 system to the 1980 system, and in survey year 2003, from the census 1990 system to the 2000 system. Of these two reclassifications, the latter was more substantial overall, yet it does not affect our sectors. The Census Bureau provides industrial crosswalks for the 1970–1980 and 1990–2000 systems, from which one can gauge how some industries are split or merged into others (Census Bureau (1989), Census Bureau (2003)). These crosswalks are basically a transition matrix for all industries from one classification to the other. A close examination of these transition "probabilities" leads us to conclude that our industries are consistently defined throughout our sample. In the transition from the 1970 system to the 1980 system, 99.9% remain inside each industry, and for the transition from the 1990 system to the 2000 system, over 95% of workers remain inside each industry. This is due to the fact that the functions of our three industries are narrowly and well defined, and are not too large.

Education and experience

We define five educational categories: "less than 12 years of schooling", "high school graduate", "13–15 years of schooling", "college graduate" (four-year college), and "more than college" (graduate degrees such as a JD, MBA, or PhD). Until survey year 1991 years of education are reported in annual steps, starting with zero through 18 years (which also absorbs instances of more than 18 years). In addition, until survey year 1991 we correct years of schooling for individuals who did not complete the last year in school by subtracting one year. This correction is not needed after survey year 1992. From survey year 1992 on early school attainment is lumped into groups: zero years, one to four years, five to six years, and seven to eight years. Also starting in survey year 1992, school attainment starting with high school is marked by degrees, not years; therefore it is not possible to distinguish between, for example, 13, 14, and 15 years of school. To make our education variable consistent throughout our sample, we adopt the coding that starts in survey year 1992, that is, we group early school attainment into brackets for the entire sample and assign maximal values to each bracket. We also group 13, 14, and 15 years of school together and assign 14 years for all individuals within that bracket in all years. In addition, we lump 17 years of schooling together with 16 years, for similar reasons. This makes the educational shares smooth throughout the sample, particularly around the 1991–1992 surveys. Experience is potential labor market experience. It is measured as $min\{age - edu - 6; age - 18\}$, where edu is years of schooling. The CPS does not contain data on job spells.

Wages and top-coding

We use the wage and salary income as our measure of wages. We deflate all wages using the deflator for personal consumption expenditures from the BEA. The reference year is 2000. Hourly wages are calculated by dividing the annual wage income by the number of hours worked. The CPS underestimates the income of individuals who earn very high salaries due to top-coding of income. Therefore the wages that we report may not be accurate for certain occupations, particularly concerning brokers, and securities and financial asset sales. In our sample the percentage of top-coded observations in the private sector increases from 0.06% in 1967 to 1.1% in 1980, after which it fluctuates in the range 0.38–1.6%, due to secular adjustments of the top-coding income limit. However the incidence of top-coding in finance relative to the nonfarm private sector is higher: In credit intermediation there are, on average, twice as many top-coded observations; in insurance there are, on average, 2.4 as many top-coded observations; and in other finance there are, on average, 13 times as many top-coded observations. This leads to an under-estimation of relative wages in the financial sector. In an attempt to compensate for this, we multiply top-coded incomes in all survey years until 1995 by a factor of 1.75. From survey years 1996 on, top-coded incomes are the average amounts of actual earnings for 12 socioeconomic cells; therefore we do not adjust them.

C.2 Matched CPS

We thank Donghoon Lee for providing us with his methodology. The matched CPS takes advantage of the fact that households in the CPS are sampled for more than a year, in the following pattern. Each household that enters the survey at any given month is sampled for four months, leaves for eight months, and then returns for four more months, after which it exits. Therefore, theoretically, every household that is surveyed in March of any given year must have been surveyed the previous March or will be surveyed the next. Of course, in practice not all individuals get surveyed twice due to survey attrition, non-compliance, and so forth.

Unfortunately, the CPS does not hold a definitive person identifier by which one could easily match two observations on the same individual from two consecutive surveys. The following methodology is used. We match individual observations from two consecutive surveys by household ID, their "line" within the household (an intra-household identifier), state of residence, race, sex, and year of birth. These are supplemented with a few more identifiers generated by the CPS (segment number, serial number, and a random cluster code). We make sure that there are only two observations within each cell defined by these identifiers and drop all the others.

Some survey years cannot be matched. Survey year 1968 cannot be matched backward because our sample starts with that year. Likewise, survey year 2006 cannot be matched forward because our sample ends with that year. Other survey years that cannot be matched for technical reasons are 1971, 1972, 1976, 1985, 1995, and 2001. Approximately 93% of all observations are actually matched from within survey years that can be matched. One important reason that we do not achieve perfect matching is that individuals who change their residential address are dropped from the sample. This affects mostly young people, but also job switchers, who may decide to move on account of changing jobs.

Definition of unemployment

Here we give the exact definition of our unemployment indicator. A person would get a positive indication of unemployment for any of the following reasons:

1. He or she did not work last year and reported being unable to find work, looking for work, or on layoff.
2. In survey years 1968–1993 the major activity in the week before the survey was looking for work.
3. In survey years 1968–1993 he or she did not work last week due to being laid-off.
4. In survey years 1994–2006 he or she reported being on layoff or looking for work.
5. In survey years 1968–1988 he or she reported the reason for working only part of the year was to look for work or due to being unemployed.
6. He or she reported a positive number of weeks looking for work last year.
7. He or she reported a positive number of weeks being unemployed last year.

Since the sample for our transition regressions includes only people that were not unemployed in the first year they were surveyed, this eventually reduces our sample.

D Imputing education shares for 1910–1930

For the period 1910–1930, when schooling data are not available, we impute the share of employees with more than high school education by occupation, separately for each sector (the nonfarm private sector and the financial sector). Although occupational classifications change across censuses, the Integrated Public Use Microdata Series (IPUMS) provides a consistent classification for occupations that is based on the 1950 census. Essentially, occupational classifications from other years are matched with the classification of 1950.

We calculate the share of workers with more than high school education in each occupation c separately for each sector j according to this classification in 1950, $\alpha_{c,j}^{1950}$. We use 1950 as a base year rather than 1940 because 1950 contains all possible occupations according to this classification, whereas 1940 is missing several. We use $\alpha_{c,j}^{1950}$ as a base to impute the share in each sector in 1910–1930 by using the distribution across occupations in each sector, $\lambda_{c,j}^{t}$, and then aggregating up:

$$s_{j,t}^{e,imputed} = \sum_{c} \lambda_{c,j}^{t} \alpha_{c,j}^{1950} \ ,$$

where $t = 1910, 1920, 1930$, $\lambda_{c,j}^{t} = \sum_{i \in c} \omega_{i,j,t} \, / \sum_{i} \omega_{i,j,t}$ is the share of workers in occupation c in sector j in census t, and $\omega_{i,t}$ is the sampling weight for that observation.

We compare the imputations to the data when possible. Figure A1 displays $\rho_{fin,t}^{imputed} \equiv s_{fin,t}^{e,imputed} - s_{nonfarm,t}^{e,imputed}$ based on 1940 and 1950, using the data in 1940–2000. In addition to shares of workers with more than high school education, we display shares of high school graduates and college graduates.

The evolution of $\rho^{imputed}_{fin,t}$ follows very similar patterns as the data in the decades near the base year, especially for shares of workers with more than high school education and college graduates. Differences in backward imputation arising from using 1940 as the base year versus 1950 are negligible. As the educational composition of occupations changes dramatically, especially after 1970, $\rho^{imputed}_{fin,t}$ underestimates $\rho_{fin,t}$. However, the timing in the changes in the trends is very similar to that in the data. This leads us to believe that if there is a bias in the backward imputation, then it is likely to be downward, not upward. If the 1990s and 2000s are similar to the 1920s and 1930s, then this is what we would expect.

E Top wage series

Using the BEA data, the average full-time equivalent wage in sector s at time t is

$$w_{s,t} = \frac{NW_{s,t}}{FTE_{s,t}} = \frac{NW_{s,t}}{N_{s,t}}\theta_{s,t} \ ,$$

where NW is the compensation of employees, FTE is full-time equivalent employment, N is employment, and θ is the hours correction, as described above. This can be written as

$$\begin{aligned} w_{s,t} &= \frac{\sum\limits_{i\in s} w_{i,t}}{\sum\limits_{i\in s} 1}\cdot\theta_{s,t} = \left[\frac{\sum\limits_{i\in\langle s,b\rangle} w_{i,t} + \sum\limits_{i\in\langle s,top\rangle} w_{i,t}}{\sum\limits_{i\in s} 1}\right]\cdot\theta_{s,t} \\ &= \left[\frac{\sum\limits_{i\in\langle s,b\rangle} 1}{\sum\limits_{i\in s} 1}\frac{\sum\limits_{i\in\langle s,b\rangle} w_{i,t}}{\sum\limits_{i\in\langle s,b\rangle} 1} + \frac{\sum\limits_{i\in\langle s,top\rangle} 1}{\sum\limits_{i\in s} 1}\frac{\sum\limits_{i\in\langle s,top\rangle} w_{i,t}}{\sum\limits_{i\in\langle s,top\rangle} 1}\right]\cdot\theta_{s,t} = \left[\alpha^b_{s,t}w^b_{s,t} + \left(1-\alpha^b_{s,t}\right)w^{top}_{s,t}\right]\cdot\theta_{s,t} \ , \end{aligned}$$

where $w^b_{s,t}$ is the average wage of the bottom $\alpha^b_{s,t}$ workers in terms of wages ($0 < \alpha^b_{s,t} < 1$), and $w^{top}_{s,t}$ is the average wage for the remaining high wage workers. Here $i \in \langle s,b\rangle$ indicates that individual i earns no more than the $\alpha^b_{s,t}$ percentile wage. Similarly, $i \in \langle s,top\rangle$ indicates that individual i earns strictly more than the $\alpha^b_{s,t}$ percentile wage. We take into account sampling weights when calculating $w^b_{s,t}$: $w^b_{s,t} = \sum\limits_{i\in s} w_{i,t}\lambda_{i,t}b_{i,t}/\sum\limits_{i\in s}\lambda_{i,t}b_{i,t}$ and $b_{i,t}$ is an indicator for individual i earning no more than the $\alpha^b_{s,t}$ percentile wage.

Thus we can write the average wage in sector s at time t as

$$w_{s,t} = \left[\alpha^b_{s,t}w^b_{s,t} + \left(1-\alpha^b_{s,t}\right)w^{top}_{s,t}\right]\cdot\theta_{s,t} \ ,$$

where $w^b_{s,t}$ is the average wage of the bottom $\alpha^b_{s,t}$ workers in the wage distribution ($0 < \alpha^b_{s,t} < 1$) and $w^{top}_{s,t}$ is the average wage for the remaining workers. The factor $\theta_{s,t}$ corrects for full-time equivalents. Since $w_{s,t}$ already takes this into account in the BEA data, it does not appear on the left hand side of the equation. Rearranging, we calculate the average full-time equivalent wage:

$$w^{top}_{s,t}\theta_{s,t} = \frac{w_{s,t} - \alpha^b_{s,t}w^b_{s,t}\theta_{s,t}}{1-\alpha^b_{s,t}} \ .$$

We obtain $\theta_{s,t}$ by dividing aggregate employment by full-time equivalent employment by sector (both series from the BEA Industry Accounts) and $w_{s,t}$ is described above. We compute $\alpha^b_{s,t}$ and $w^b_{s,t}$ from the 1940–2000 censuses and the 2010 ACS. In doing so, we take into account hours worked and sampling weights. We use $\alpha^b_{s,t} = 0.90$ so that $w^{top}_{s,t}\theta_{s,t}$ is the average wage of the top decile in full-time equivalent terms. This is what is used in the main text and plotted in Figure II. Figure A2 reports the $w^{top}_{s,t}\theta_{s,t}$ series using $\alpha^b_{s,t} = 0.75$ and $\alpha^b_{s,t} = 0.5$, so that we consider the top quartile wage and median wage.

F Relative wage decompositions

We decompose changes in ω_{fin} into within and between group changes along several dimensions using the equation from the main text. We apply this decomposition in three subsamples: 1933–1960, 1960–1980, and 1980–2005. The year 1933 marks the beginning of the regulated period in finance; 1960 marks the beginning of the most regulated period; 1980 marks the beginning of the deregulation period. Data availability constrain some decompositions to start in 1969. The decompositions are reported in Table II.

Total relative wage changes using the CPS are smaller than when using the BEA Industry Accounts. The reason is top-coding in the CPS. Thus, while in the BEA Industry Accounts the relative wage of finance increases by 0.65 from 1980 to 2005, in the CPS it increases only by 0.4. Therefore the wages that we report may not be accurate for certain occupations, particularly brokers and traders.

Educational and occupational decompositions

We use data from the CPS and compute the relative wage of finance, using total annual hours worked and total income from wage and salary in finance and the nonfarm private sector. We decompose the relative wage of finance into five educational categories: "less than 12 years of schooling", "high school graduate", "13–15 years of schooling", "college graduate" (four-year college), and "more than college" (graduate degrees such as a JD, MBA, or PhD). We then decompose across seven occupational groups: "managers and professionals", "mathematics and computers", "insurance specialists", "brokers and traders", "bank tellers", "administration, including clerks" and "all the rest" (janitors, security and miscellaneous). Our classification of occupations attempts to group employees according to the tasks they perform.

Sectoral decompositions

We first decompose the finance relative wage along subsectors within finance: credit intermediation, insurance, and other finance. We then decompose changes in $\omega_{nonfarm} = 1/\omega_{fin}$ across 11 subsectors within the nonfarm private sector, excluding finance: mining; construction; manufacturing; transportation + utilities + broadcasting + telecommunications; wholesale trade; retail trade + food services; real estate; legal services; educational services, health + social services; and other services. Our choices of the groupings of industries are done to achieve comparability across classifications over time. We use the same decomposition formula, except that now the signs are reversed because we are decomposing the inverse of ω_{fin}. All series are computed in full-time equivalent terms, using data from the BEA Industry Accounts.

Geographical decomposition

We decompose the finance relative wage along 50 states + Washington DC. Table A1 reports the geographical decomposition in much more detail. If finance becomes more concentrated over time in high-earning financial centers, then the between component should dominate. However only the within component matters: Changes in the geographical distribution do not matter much. The Herfindahl index for the wage bill share of finance across states is stable, whereas the Herfindahl index for employment shares falls in the sample. These numbers are available upon request.

G Relative task intensity indices

In order to construct our relative task intensity indices we match the occupational task intensity indices from the Dictionary of Occupational Titles (DOT) into individual occupations in the censuses from 1910 to 2000 and in the 2008 March CPS. Five DOT task intensities by occupation and gender (373×2 cells for each task type) were obtained from David Autor, to whom we are grateful for sharing these data. The occupations are classified according to the 1990 census system. The task intensity measures vary over the [0,10] interval. We call these data DOT1990. Census and CPS data are extracted from IPUMS.

DOT task intensities

The DOT task intensities were originally calculated in 1977 by a panel of experts from the National Academy of Sciences for 3,886 DOT occupations. Each occupation was assigned a vector of characteristics. From this vector we use only five elements that sufficiently characterize each occupation: finger dexterity (routine manual tasks); set limits, tolerances, and standards (routine cognitive tasks); math aptitude (analytical thinking); direction, control, and planning (decision making); and eye–hand–foot coordination (captures non-routine manual tasks). The 3,886 DOT occupations are allocated across 411 occupations of the 1970 census classification. The task intensity for each 1970 census occupation is a weighted average over the tasks of the original DOT occupations allocated to it, where the weights are CPS sampling weights. This was achieved using the April 1971 CPS (which pertains to 1970). The averages are different for men and women, hence the separation by gender. Each one of the five indices is detected as a principal component for indices that are similar in nature (see Autor, Levy, and Murnane (2003)). The 1970 census classification is matched into the 1990 census classification, using information based on the OCC1990 variable in IPUMS (this was carried out by Peter Meyer from the Bureau of Labor Statistics).

Consistent occupational classification

To match the DOT1990 data to occupations in 1910–2007 we had to create a consistent classification system for the entire period. For 1960–2007 we can use the 1990 census classification directly, using the OCC1990 variable in IPUMS. For 1910–1950 we use the 1950 census classification, using the OCC1950 variable in IPUMS. We create a crosswalk for OCC1950 into OCC1990 using the 1950 census, the first year for which OCC1990 exists. We used 1950 as a base for the crosswalk because all census 1950 occupations appear in 1950. Another option we tried was to use the 1990 census as the base for the crosswalk; this had no effect on our results.

When matching the DOT1990 data we make a few modifications due to the fact that not all of the 1990 census occupations are represented in DOT1990. Therefore we allocate task intensities to these occupations using data for other occupations that we think are very similar in nature, a priori. The only substantial modification is to allocate task intensities to "professionals, not elsewhere classified" according to the average task intensity for professionals by year, 2-digit industry, and gender. Our results are not affected by dropping all the occupations that are not matched or to modifications of these allocations.

Eventually, we construct a data set with a consistent classification of occupations. The DOT1990 information is then matched into this data set, using the 1990 census classification and gender. Thus every individual in the data set has five task intensity indices that characterize his or her occupation.

Aggregation

We restrict attention to workers aged 15 to 65 who are employed in the nonfarm private sector (in 1920 we can only restrict to individuals who are in the labor force). For each task and year we aggregate up by sector as follows:

$$task_{s,t} = \frac{\sum_{i \in s} task_i \ \lambda_{i,t} \ hrs_{i,t}}{\sum_{i \in s} \lambda_{i,t} \ hrs_{i,t}} \ ,$$

where i denotes a particular individual, t denotes the year, the λ are sampling weights, and the hrs are annual hours. Here $i \in s$ means that individual i works in sector s, where $s = fin$ corresponds to the financial sector and $s = nonfarm$ corresponds to the nonfarm private sector. The generic $task$ varies over all five tasks described above.

Unfortunately, it is not possible to calculate hrs for all years. In the 1910–1930 and 1960–1970 censuses the underlying data to do so are missing. Therefore in those years we treat $hrs = 1$ for all individuals. The underlying data used to calculate hrs is the number of weeks worked times the number of hours worked per week. The 1910–1930 censuses do not contain any such information. In 1940–1950 we use data on hours worked in the week before the census. The 1960–1970 censuses contain only categorical data on weeks and hours worked according to some ad hoc intervals; we could not calculate hours worked because we could not accurately adjust for longer hours or more weeks. In the 1980–2000 censuses, as well as the 2008 March CPS, we use data on usual hours worked per week. Our attempts to gauge hours and weeks worked in 1960–1970 by using data from 1950, 1980, or both, result in severe jumps in the $task$ series in those years.

The relative task intensity for finance for each year is given by

$$rel\_task_{fin,t} \equiv task_{fin,t} - task_{nonfarm,t} \ .$$

H Executive compensation

We obtain data from Frydman and Saks (2010) on executive compensation in 50 of the largest publicly traded firms operating in the U.S. in 1936–2005. We thank Carola Frydman and Raven Saks for making these data available to us. These firms report executive compensation for at least 20 years within at least one of three windows (1936–1966, 1943–1973, and 1970–2000). Frydman and Saks (2010) argue that their sample is representative of the top 300 firms in the U.S. during 1936–2005. They also show that the trends in their data are comparable to other sources, such as Execucomp and *Forbes Magazine*, when the data are available. Of these 50 firms, in any given year seven are included in the financial sector; none are in agriculture. None of the financial firms in the sample span the entire period. The coverage is as follows: CIT Group 1938–1976, Citicorp (Citigroup) 1971–1997, American Express 1977–2005, Chase (JPMorgan Chase) 1972–2005, Aetna 1964–2005, Cigna 1982–2005, and AIG 1970–2005.

Each firm reports compensation for the top three officers (not just the chief executive officer, or CEO) in 10-K reports (1936–1941), in proxy statements (1942–1991), or in Compustat (1992–2005). Compensation includes salary, bonus, and option value. Most bonuses are paid in cash. Bonuses that are paid in stock are evaluated using the stock price at the time they are granted. The value of options at the time they are granted is calculated using the Black–Scholes formula. For full documentation we refer the reader to Frydman and Saks (2010). We denote the median compensation for the top three executives outside of finance by $wage^{exec}_{nonfarm,t}$ and that in finance by $wage^{exec}_{fin,t}$. We do not find jumps or discontinuities in the $wage^{exec}_{fin,t}$ series around the years in which a financial firm joins or leave the sample.

We obtain additional data on CEO compensation from Execucomp and merge it into the Compustat data for 1992–2010. For this sample we start with the 1,000 top firms by CEO compensation before taxes and drop firms in agriculture and firms with no market capitalization information. This leaves us with 809 firms per year, on average, with stable representation of financial firms at 9.8%, on average. Only in 1992 and 1993 is the representation of finance lower, at 4.5% and 6.75%, respectively. We convert nominal values into 2010 prices using the Consumer Price Index series from the Bureau of Labor Statistics. We use these data to estimate dual scaling models as in Gabaix and Landier (2008).

We obtain data on CEO compensation collected by Kevin J. Murphy from *Forbes Magazine* in 1970–1991; we thank him for sharing these data with us. These data do not include information on options until 1978. After that, they include gains from exercising options (and not grant date values, as in Frydman and Saks (2010)).

We obtain data on market capitalization for 1927–2010 from the Center for Research in Securities Prices (CRSP). Again, we drop firms in agriculture and firms with no market capitalization information. Over time the number of firms, as well as the representation of financial firms, increase. We convert nominal values into 2010 prices using the Consumer Price Index series. We use this data to estimate relative executive compensation in finance. The estimation relies on predictions of the estimated relation from in 1992–2010.

Relative market value for Frydman and Saks (2010) sample

We construct the relative median market size of finance for the sample of Frydman and Saks (2010), using data from CRSP. To be consistent with the data used to construct $\omega^{exec}_{fin,t}$ we follow the sampling methodology of Frydman and Saks (2010): We restrict attention to 50 firms who have reported for at least 20 years; the number of financial firms in each year follows the sample above (one firm in 1938–1963, two in 1964–1969, three in 1970, four in 1971, five in 1972–1981 and 1998–2005, and six in 1982–1997). All firms are the largest in their subsector (finance or nonfarm private sector). The relative median market size in this sample is reported in Figure A4, together with the executive compensation series used in Figure IV, to facilitate comparison. The relative market value is lower than relative executive compensation. The relative median market value increases from 0.45 in 1966–1975 to 1.43 in 1996–2005, or 1.86 in 2001–2005 (on average), an increase of 3.18 or 4.13 times, respectively.

Comparing data across sources in 1970–2005

We compare the relative executive compensation series based on Frydman and Saks (2010) with the relative CEO compensation series constructed by combining the *Forbes* data with that of Execucomp in 1970–2005. We first compute the ratio of the median, top quartile, and top decile CEO compensation in finance to the respective percentiles in the rest of the nonfarm private sector from the entire *Forbes*–Execucomp sample; see Figure A3-A. To get closer to the sampling of Frydman and Saks (2010) we also consider the ratio of the median CEO compensation for the five highest paid CEOs in finance to the median for the 45 highest paid CEOs in the nonfarm private sector; see Figure A3-B. Note that this is not the same as the 50 largest firms. The patterns are similar to those in Figure IV in the relevant years.

Predictions for segregated labor markets for executives

We consider the case where the labor markets for executives in finance and elsewhere are segregated. We do this in two ways: First, we estimate the dual scaling model separately for finance and elsewhere, using data from Execucomp and Compustat (see the Appendix for details). Table A4, columns 4–9, reports the results. The coefficients are much smaller in finance, both economically and statistically. The predictions of $\omega^{exec}_{fin,t}$ based on these estimates are nonsensical: They imply a sharp drop in the relative executive compensation in finance from the mid-1930s to the mid-1960s, a small recovery up to the mid-1980s, and then a steady decline. This is true both when using the sampling methodology of Frydman and Saks (2010) or when restricting the CRSP data to the largest 400 firms in any year. These results are available upon request.

Another way to consider segregated labor markets for executives is to assume that the mechanics of the assignment of executives to firms is identical in both sectors, that is, α and β are the same in both sectors. In the model of Gabaix and Landier (2008) α is the Pareto coefficient for the firm size distribution; we estimate this in the CRSP data and find it to be similar (and close to unity) in finance and the nonfarm private sector for "normal", non-crisis years. Here β is the Pareto coefficient for the executive talent distribution (an approximation for the tail of the talent distribution); unfortunately, it is not possible to study whether this parameter varies across sectors without estimating the model completely. Therefore we allow only the return to scale parameter γ and the constants to vary across subsectors in the dual scaling model. This allows only the idiosyncratic slopes to vary. We also allow the reference firm (median firm) to vary across sectors. Table A4 reports the results in columns 10–12.

The predictions of changes in $\omega^{exec}_{fin,t}$ based on these estimates are similar to those using the integrated markets approach in the main text, although the levels are completely off. Using the sampling methodology of Frydman and Saks (2010), the predicted relative compensation falls from 0.30 in 1935 to 0.24 in 1966–1975 and then increases to 0.57 in 1996–2005 (on average)—an increase of 2.4, which is still much smaller than the 3.5 increase in $\omega^{exec}_{fin,t}$ in the data. Figure A6 reports the results graphically, where we adjust the benchmark to be equal to the average $\omega^{exec}_{fin,t}$ in the data in 1966–1975. The results of Figure A6 are virtually indistinguishable from those in Figures X-C and X-D

I Inequality simulation

We use the sample of workers in finance in 1970, denoted $FI70$, as a base to simulate wages in finance in all the other years. We define this sample as $\{\lambda_i, w_i, X_i\}_{i\in FI70}$, where the λ are the CPS sampling weights, the w are annual wages, and X is a vector of characteristics (to be used for calculating residual inequality). In all the other years $t = 1967$ to 2005 observations in finance are simulated as $\{\lambda_i \cdot \kappa_t, w_i \cdot (1+\gamma_t), X_i\}_{i\in FI70}$, where $\kappa_t = \left(\sum_{i\in fi} \lambda_{it}\right) / \left(\sum_{i\in FI70} \lambda_i\right)$ updates sampling weights to keep the same sum of weights as in the original data and γ_t denotes the growth of the median wage relative to 1970. To fix employment shares we further multiply sampling weights in finance by a factor of s^{fi}_{1970}/s^{fi}_t and those in the rest of the private sector by s^{ps}_{1970}/s^{ps}_t, where $s^{fi}_t = \left(\sum_{i\in fi} \lambda_{it}\right) / \left(\sum_i \lambda_{it}\right)$ is the employment share of finance in year t, and similarly for the nonfarm private sector (ps). Updating sampling weights is important because the measures of inequality take these weights into account directly. For example, percentiles are calculated according to the weighted position in the distribution. The median wage in some year is given by w_j such that j solves $\left(\sum_{i\leq j} \lambda_i\right) / \left(\sum_i \lambda_i\right) = 0.5$, where the observations are arranged in ascending order of wages. In addition, updating weights is a natural way to update the number of people across years.

The sample in which wages in finance are replaced by simulated wages as described above is called the simulated sample.

J Explanatory variables for regressions

Information technology

We compute the share of IT and software in the capital stock in each sector as follows. We use data from the BEA's fixed assets tables by industry, which provide both current-cost net capital stock of private nonresidential fixed assets, as well as chain-type quantity indices for these fixed assets. Denote k_t as current-cost net capital stock of some fixed asset, and denote q^{2000}_t as chain-type quantity index for that asset, where $q^{2000}_{2000} = 100$. We use the following formula to get constant price values for each fixed asset, $k_t^{2000} = q_t^{2000} \cdot k_{2000}/100$. The IT intensity variable is computed as the share of the k_t^{2000} series for computers and peripheral equipment and software divided by the k_t^{2000} series for the aggregate equipment fixed assets.

Use of patents in finance
We use data on new patents used by industry in 1909–1996 from Carter, Gartner, Haines, Olmstead, Sutch, and Wright (2006). These data are based on methodology developed by Daniel Johnson at Colorado College. We normalize the number of patents used in finance by the total number of patents used. We use data on financial patent counts from Lerner (2006) to extrapolate the normalized patent use series to 2002.

Corporate finance activity: initial public offerings and credit risk
We measure IPO activity from 1900 to 2002 by dividing the market value of IPOs by the market value of existing equities, using data from Jovanovic and Rousseau (2005). We measure credit risk by a three-year moving average of the U.S. corporate default rate published by Moody's. We normalize the IPO and credit risk series to have a mean of zero and unit standard deviation over the sample period.

Financial deregulation index
We construct a measure of financial deregulation that takes into account branching restrictions, the Glass–Steagall Act, interest ceilings, the separation of insurance companies from banks, and restrictions on the investment opportunities of insurance companies and banks.

(i) Branching
We use the share of the U.S. population living in states that have removed branching restrictions via mergers and acquisitions. The data are from Black and Strahan (2001). Our branching deregulation indicator is a continuous variable that starts at 16.7% in 1960 and increases to 100% by 1999. We set our indicator at 16.7% from 1927 to 1960. The McFadden Act of 1927 prevents branching of nationally chartered banks; before the McFadden Act branching was less clearly limited. To capture this, we set our indicator to 0.3 in the years 1909–1926.

(ii) Separation of commercial and investment banks
The Glass-Steagall indicator is a continuous variable between 0 and 1. It is 0 until 1932, 0.5 in 1933 and 1 from 1934 to 1986. The Glass-Steagall act is relaxed in 1987, 1989, 1997 and was finally repealed in 1999, by the Gramm-Leach-Bliley Act. In 2000 this indicator is back to zero.

(iii) Interest rates ceilings
Ceilings were introduced in 1933 and removed after 1980. Our indicator variable is zero until 1932, 0.5 in 1933, and one from 1934 to 1980. Saving and loans institutions were further deregulated by the Garn–St. Germain Depository Institutions Act of 1982. To capture these features, our index moves gradually to zero between 1980 and 1983.

(iv) Separation of banks and insurance companies
The Bank Holding Company Act of 1956 prohibited a bank holding company from engaging in most non-banking activities and from acquiring voting securities of certain companies. It was repealed in 1999. The Armstrong investigation of 1905 took place before the beginning of our sample and therefore is not directly relevant.

The deregulation index is given by

$$deregulation = (i) - (ii) - (iii) - (iv) \ .$$

Financial globalization
We proxy for external demand forces such as financial globalization by using the ratio of U.S. foreign assets to GDP. The data on foreign assets are from Obstfeld and Taylor (2004) (1900–1960) and the International Monetary Fund (1980–2005). We interpolate linearly between data points when data are missing.

Top marginal tax rate
The top marginal tax rate controls for either the supply of talented individuals or for the cost of paying high net wages. Tax rate data are from the Tax Foundation, based on information from the U.S. Internal Revenue Service.

K Unemployment risk calibration

Philippon and Reshef (2007) find that unemployment risk in finance increases relative to the nonfarm private sector by 2.5 percentage points from 1971–1980 to 1991–2005. We use this finding to estimate the compensating differential in wages that is required to keep workers indifferent to this increase in risk.

Ruhm (1991) finds that layoffs lead to temporary unemployment and long-lasting decreases in earnings: "Displaced workers were out of work eight weeks more than their observably similar counterparts in the year of the separation, four additional weeks in period $t+1$, and two extra weeks at $t+2$. By year $t+3$ they were jobless only 1.5 weeks more than the peer group, and the $t+4$ increase was just six days". By contrast, "almost none of the $t+1$ wage reduction dissipated with time. The earnings gap remained at 13.8 percent and 13.7 percent, respectively, in years $t+3$ and $t+4$."

A complete study of the effects of unemployment risk on the level of compensation that is needed to keep workers indifferent between different jobs is clearly beyond the scope of this paper. Nonetheless, we think it is useful to provide some simple benchmark calculations. We do so in the simplest framework possible and assume that labor income is the only source of risk and that the utility function has constant relative risk aversion. We set the personal discount rate and market rate both equal to 3% per year. We assume that workers live and work for 40 years and that the labor income process, y_t, is given by

$$y_{t+1} = \left\{ \begin{array}{c} 1.02\ y_t \text{ with probability } 1-p \\ 0.9\ y_t \text{ with probability } p \end{array} \right\},$$

with y_1 given. The increase of 2% captures the normal increase in real labor income. The drop by 10% captures the income loss from displacement documented by Ruhm (1991). This process implies that shocks are permanent, which makes the effect of unemployment risk more important, so we are likely to obtain an upper bound for the impact on relative wages.

We perform the following experiment. First we set $p = 4.41\%$ and $y_1 = 1$ and solve and simulate the model with a coefficient of relative risk aversion equal to two. We then increase the unemployment risk to $p = 6.91\%$. This increase of 2.5 percentage points corresponds to the increase in the relative unemployment risk documented in Philippon and Reshef (2007). To keep workers indifferent, the new starting wage should be $y_1 = 1.063$, an increase of 6%. If we lower the calibrated risk aversion to one, the required increase in wages is 6%. If we increase risk aversion to three, the required increase in wages is 6.6%.

References

Autor, D. H., F. Levy, and R. J. Murnane (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 118(4), 1279–1333.

Black, S. E., and P. E. Strahan (2001): "The Division of Spoils: Rent-Sharing and Discrimination in a Regulated Industry," *American Economic Review*, 91(4), 814–831.

Bureau of Economic Analysis (2011a): "Measuring the Nation's Economy: An Industry Perspective, A Primer on BEA's Industry Accounts," Technical Paper.

——— (2011b): "State Personal Income and Employment Methodology," Technical Paper.

Carter, S., S. Gartner, M. Haines, A. Olmstead, R. Sutch, and G. Wright (eds.) (2006): *Historical Statistics of the United States Millennial Edition Online*Cambridge University Press.

Census Bureau (1989): "The Relationship Between The 1970 and 1980 Industry and Occupation Classification Systems," Technical Paper 59.

——— (2003): "The Relationship Between the 1990 Census and Census 2000 Industry and Occupation Classification Systems," Technical Paper 65.

Frydman, C., and R. Saks (2010): "Executive Compensation: A New View from a Long-Run Perspective, 1936-2005," *The Review of Financial Studies*, 23(5), 2099–2138.

Gabaix, X., and A. Landier (2008): "Why Has CEO Pay Increased So Much?," *The Quarterly Journal of Economics*, 123(1), 49–100.

Jovanovic, B., and P. L. Rousseau (2005): "General Purpose Technologies," in *Handbook of Economic Growth*, ed. by P. Aghion, and S. Durlauf. Elsevier, Chapter 18.

Kuznets, S. (1941): "National Income and Its Composition, 1919-1938," Discussion paper, National Bureau of Economic Research.

Lerner, J. (2006): "The New New Financial Thing: The Origins of Financial Innovations," *Journal of Financial Economics*, 79(2), 223–255.

Martin, R. F. (1939): *National Income in the United States, 1799-1938*. National Industrial Conference Board,.

Mitchell, W. e. A. (1921): "Income in the United States 1909-1919," Discussion paper, National Bureau of Economic Research.

Obstfeld, M., and A. M. Taylor (2004): *Global Capital Markets: Integration, Crisis and Growth*. Cambridge University Press, New York.

Philippon, T., and A. Reshef (2007): "Skill Biased Financial Development: Education, Wages and Occupations in the U.S. Financial Sector," NBER Working Paper 13437.

Ruhm, C. J. (1991): "Are Workers Permanently Scarred by Job Displacements?," *The American Economic Review*, 81(1), 319–324.

Table A1: Decomposition of Relative Wage of Finance Across States: 1980-2005

State	(1) Average Employment Share	(2) Change in Relative Wage	(3) Average Relative Wage	(4) Change in Employment Share	(5) Within (=1*2)	(6) Between (=3*4)	(7) Contribution of (5) to Total Change (%)	(8) Contribution of (5)+(6) to Total Change (%)
New York	0.103	1.914	2.231	-0.034	0.197	-0.077	26.1%	15.9%
Connecticut	0.020	1.184	1.599	0.000	0.024	-0.001	3.2%	3.1%
Massachusetts	0.030	0.938	1.454	-0.003	0.028	-0.005	3.7%	3.0%
Illinois	0.057	0.906	1.423	-0.008	0.052	-0.011	6.8%	5.4%
Vermont	0.002	0.586	1.383	0.000	0.001	0.000	0.1%	0.1%
Maine	0.004	0.517	1.350	0.001	0.002	0.001	0.2%	0.4%
California	0.115	0.777	1.342	-0.001	0.089	-0.001	11.8%	11.6%
Minnesota	0.021	0.764	1.325	0.003	0.016	0.004	2.1%	2.7%
Rhode Island	0.004	0.559	1.315	0.000	0.002	0.000	0.3%	0.3%
Florida	0.053	0.579	1.310	0.018	0.030	0.024	4.0%	7.2%
North Carolina	0.022	0.827	1.301	0.003	0.018	0.004	2.4%	3.0%
Virginia	0.022	1.038	1.300	0.000	0.022	0.000	3.0%	2.9%
Maryland	0.016	0.613	1.299	0.004	0.010	0.005	1.3%	2.0%
New Jersey	0.033	0.749	1.287	0.005	0.024	0.006	3.2%	4.0%
Iowa	0.013	0.762	1.274	0.000	0.010	0.000	1.3%	1.3%
New Hampshire	0.004	0.849	1.262	0.000	0.004	0.000	0.5%	0.5%
Nebraska	0.008	0.426	1.241	0.000	0.003	0.000	0.5%	0.5%
Georgia	0.025	0.700	1.225	0.002	0.018	0.003	2.4%	2.7%
Alabama	0.012	0.616	1.220	0.000	0.007	-0.001	1.0%	0.9%
Pennsylvania	0.045	0.661	1.198	-0.005	0.030	-0.006	3.9%	3.1%
Washington	0.017	0.460	1.194	0.003	0.008	0.003	1.0%	1.5%
South Carolina	0.010	0.339	1.186	0.003	0.003	0.003	0.4%	0.8%
South Dakota	0.003	0.266	1.183	0.001	0.001	0.002	0.1%	0.3%
Tennessee	0.018	0.717	1.179	-0.002	0.013	-0.002	1.7%	1.5%
Missouri	0.020	0.503	1.160	-0.002	0.010	-0.002	1.4%	1.0%
Oregon	0.011	0.594	1.157	-0.001	0.006	-0.001	0.8%	0.7%
Kentucky	0.011	0.698	1.151	-0.001	0.008	-0.001	1.0%	0.9%
Mississippi	0.006	0.293	1.150	-0.001	0.002	-0.001	0.2%	0.1%
Ohio	0.038	0.628	1.145	-0.004	0.024	-0.004	3.2%	2.6%
Montana	0.003	0.319	1.142	0.000	0.001	0.000	0.1%	0.1%
District of Columbia	0.004	0.775	1.138	-0.004	0.003	-0.004	0.4%	-0.1%
Arizona	0.016	0.436	1.124	0.010	0.007	0.011	0.9%	2.4%
Delaware	0.004	0.900	1.124	0.003	0.004	0.004	0.5%	1.0%
Idaho	0.003	0.337	1.124	0.001	0.001	0.001	0.1%	0.3%
Arkansas	0.006	0.223	1.122	0.000	0.001	0.000	0.2%	0.2%
Wisconsin	0.021	0.688	1.121	-0.001	0.015	-0.002	1.9%	1.7%
New Mexico	0.004	0.429	1.109	0.000	0.002	0.000	0.2%	0.3%
Colorado	0.018	0.478	1.106	0.005	0.008	0.006	1.1%	1.9%
Kansas	0.010	0.425	1.097	0.000	0.004	-0.001	0.6%	0.5%
Texas	0.073	0.535	1.094	0.011	0.039	0.013	5.2%	6.8%
Indiana	0.017	0.481	1.093	-0.002	0.008	-0.003	1.1%	0.8%
Nevada	0.006	0.501	1.085	0.006	0.003	0.006	0.4%	1.2%
Michigan	0.027	0.417	1.048	-0.001	0.011	-0.001	1.5%	1.3%
Louisiana	0.013	0.499	1.042	-0.005	0.006	-0.005	0.8%	0.2%
North Dakota	0.003	0.450	1.024	0.000	0.001	0.000	0.2%	0.1%
Utah	0.008	0.248	1.017	0.006	0.002	0.006	0.3%	1.0%
Oklahoma	0.011	0.362	1.009	-0.002	0.004	-0.002	0.5%	0.2%
Hawaii	0.005	0.478	0.994	-0.003	0.002	-0.003	0.3%	-0.1%
West Virginia	0.004	0.421	0.964	-0.001	0.002	-0.001	0.2%	0.1%
Wyoming	0.002	0.390	0.747	-0.001	0.001	0.000	0.1%	0.0%
Alaska	0.002	0.642	0.712	-0.002	0.001	-0.001	0.2%	0.0%
Total	1.000			0.000	0.790	-0.035	105%	100%
					0.755			

Notes. Decomposition of the increase in the relative wage in finance in 1980-2005. Column (5) reports the contribution of changes in relative wages within categories, while holding the geographical composition fixed at the average for the period. Column (6) reports the contribution of changes in the geographical composition, while holding relative wages fixed at the average for the period within states. Together, columns (5) and (6) sum up to the total change, according to the decomposition equation in the text. Column (7) reports the contribution of the within component of each state to the total increase (0.755). Column (8) reports the contribution of each state to the total increase. States are sorted by average relative wage (column 3). Source: our calculations based on the State Personal Income Tables.

Table A2: Historical Determinants of Education and Wages in Finance Industry, Disaggregated Regulation Index

A. Change Regressions

Dependent variable	Change in Relative Education, t to t+5				Change in Relative Wage, t to t+5			
Change in Glass-Steagall deregulation, t-5 to t	0.0006	0.0005	-0.0005	-0.0045	0.1982***	0.1982***	0.1468**	0.1170*
	(0.0049)	(0.0049)	(0.0028)	(0.0033)	(0.0461)	(0.0458)	(0.0590)	(0.0611)
Change in branching deregulation, t-5 to t	0.0645***	0.0629***	0.0624***	0.0600***	0.3794***	0.3958***	0.3558***	0.3327***
	(0.0055)	(0.0054)	(0.0053)	(0.0041)	(0.1426)	(0.1466)	(0.1134)	(0.1230)
Change in interest ceiling deregulation, t-5 to t	0.0104***	0.0103***	0.0094***	0.0094***	0.0200	0.0205	0.0092	-0.0012
	(0.0030)	(0.0030)	(0.0022)	(0.0025)	(0.0261)	(0.0253)	(0.0448)	(0.0425)
Change in bank-insurance separation deregulation, t-5 to t	0.0039*	0.0039*	0.0030***	0.0017*	-0.0300*	-0.0305*	-0.0671	-0.0725
	(0.0021)	(0.0022)	(0.0008)	(0.0009)	(0.0178)	(0.0173)	(0.0506)	(0.0475)
Change in Financial Patents over Total Patents, t-5 to t		0.8711				-8.5748		
		(0.8120)				(10.1506)		
Change IPO share of market capitalization, t-5 to t			0.0034***	0.0018**			0.0741***	0.0627**
			(0.0008)	(0.0008)			(0.0241)	(0.0238)
Change in Corporate Default Rate, t-5 to t			0.0010*	0.0006			0.0298*	0.0273
			(0.0006)	(0.0004)			(0.0163)	(0.0179)
Change in Foreign Assets/GDP, t-5 to t				0.0599***				0.4061
				(0.0219)				(0.2450)
Change in Top Marginal Tax Rate, t-5 to t				-0.0020				-0.0541
				(0.0043)				(0.0906)
Observations	91	91	77	77	93	93	78	78
R-squared	0.737	0.743	0.811	0.860	0.357	0.361	0.648	0.667
Sample	1910-2005	1910-2003	1920-2003	1920-2003	1909-2006	1909-2003	1920-2003	1920-2003

B. Level Regressions

Dependent variable	Relative Education				Relative Wage			
Glass-Steagall deregulation, t-5	0.0215***	0.0148***	0.0148***	0.0147***	0.2953***	0.2381**	0.2481***	0.2412***
	(0.0045)	(0.0052)	(0.0050)	(0.0054)	(0.0812)	(0.0983)	(0.0791)	(0.0714)
Branching deregulation, t-5	0.0538***	0.0460***	0.0530***	0.0532***	0.3767***	0.1243	0.2949***	0.3180***
	(0.0044)	(0.0049)	(0.0071)	(0.0076)	(0.0901)	(0.1279)	(0.1031)	(0.1033)
Interest ceiling deregulation, t-5	0.0110***	0.0115***	0.0109***	0.0106**	0.0445	0.0827	0.0597	0.0117
	(0.0029)	(0.0032)	(0.0030)	(0.0042)	(0.0693)	(0.1001)	(0.0768)	(0.0711)
Bank-insurance separation deregulation, t-5	0.0134***	0.0116***	0.0099***	0.0100***	0.1209**	0.0774**	0.0413	0.0493
	(0.0039)	(0.0028)	(0.0030)	(0.0032)	(0.0552)	(0.0297)	(0.0364)	(0.0367)
IPO share of market capitalization, t-5		0.0026*	0.0036**	0.0036**		0.0679***	0.0908***	0.0896***
		(0.0014)	(0.0016)	(0.0016)		(0.0214)	(0.0242)	(0.0228)
Corporate Default Rate, t-5		0.0023**	0.0026**	0.0025**		0.0451***	0.0512***	0.0486***
		(0.0011)	(0.0010)	(0.0010)		(0.0163)	(0.0159)	(0.0151)
Foreign Assets/GDP, t-5			-0.0225	-0.0231			-0.4996**	-0.5636**
			(0.0167)	(0.0181)			(0.2312)	(0.2433)
Top Marginal Tax Rate, t-5				-0.0011				-0.1422
				(0.0081)				(0.0936)
Observations	96	82	82	82	98	83	83	83
R-squared	0.948	0.959	0.960	0.960	0.851	0.908	0.916	0.920
Sample	1910-2005	1910-2003	1920-2003	1920-2003	1909-2006	1909-2003	1920-2003	1920-2003

Notes. Standard errors in parentheses; in Panels A and B Newey-West standard errors with 5 lags of autocorrelation. In Panel C the effect of branching deregulation on insurance is omitted due to colinearity considerations. *** p<0.01, ** p<0.05, * p<0.1. The regulation variables are coded such that higher values indicate more deregulation (looser regulation).

Table A2: Historical Determinants of Education and Wages in Finance Industry, Disaggregated Regulation Index

C. Panel of Subsectors: Credit Intermediation, Insurance and Other Finance.

Dependent variable	Effect on	Relative Education			Relative Wage		
Glass-Steagall deregulation, t-1	Other Finance	0.109***		0.0982***	1.374***		1.437***
		(0.0138)		(0.0124)	(0.176)		(0.176)
	Credit Intermediation	0.0160		0.00434	0.0703		0.136
		(0.0138)		(0.0124)	(0.176)		(0.176)
	Insurance	0.0520***		0.0427***	0.139		0.192
		(0.0117)		(0.0105)	(0.150)		(0.150)
Branching deregulation, t-1	Other Finance	-0.0473***		-0.0273*	1.266***		1.153***
		(0.0175)		(0.0160)	(0.224)		(0.227)
	Credit Intermediation	0.0411**		0.0471***	-0.0863		-0.121
		(0.0175)		(0.0156)	(0.224)		(0.221)
	Insurance	-		-	-		-
		-		-	-		-
Interest ceiling deregulation, t-1	Other Finance	0.0283**		0.00131	0.239		0.393**
		(0.0114)		(0.0114)	(0.146)		(0.161)
	Credit Intermediation	0.00895		-0.00167	0.00390		0.0641
		(0.0114)		(0.0103)	(0.146)		(0.147)
	Insurance	-0.00391		-0.0121*	0.0172		0.0639
		(0.00794)		(0.00721)	(0.102)		(0.102)
Bank-insurance separation deregulation, t-1	Other Finance	-0.0701***		-0.0654***	-0.133		-0.159
		(0.00806)		(0.00720)	(0.103)		(0.102)
	Credit Intermediation	-0.00129		-0.00541	0.0833		0.107
		(0.00806)		(0.00719)	(0.103)		(0.102)
	Insurance	0.00395		0.00341	-0.0264		-0.0234
		(0.00806)		(0.00715)	(0.103)		(0.101)
Share of IT in Capital Stock of Subsector (t-1)			0.197***	0.144***		1.679	-0.818**
			(0.0449)	(0.0273)		(1.415)	(0.387)
Subsector fixed effects		Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects		Yes	Yes	Yes	Yes	Yes	Yes
Observations		165	165	165	165	165	165
R-squared		0.914	0.704	0.933	0.973	0.443	0.975
Sample		1951-2005	1951-2005	1951-2005	1951-2005	1951-2005	1951-2005

Notes. Standard errors in parentheses; in Panels A and B Newey-West standard errors with 5 lags of autocorrelation. In Panel C the effect of branching deregulation on insurance is omitted due to colinearity considerations. *** p<0.01, ** p<0.05, * p<0.1. The regulation variables are coded such that higher values indicate more deregulation (looser regulation).

Table A3: Historical Determinants of Education and Wages in Finance Industry, Alternative Specification

Dependent variable	Relative Education				Relative Wage			
Deregulation Index	0.0175***	0.0134***			0.1489***	0.0942***		
	(0.0009)	(0.0023)			(0.0103)	(0.0175)		
Glass-Steagall deregulation			0.0224**	0.0188**			0.2040*	0.1215
			(0.0099)	(0.0091)			(0.1076)	(0.0963)
Branching deregulation			0.0303***	0.0493***			0.1927***	0.0317
			(0.0085)	(0.0119)			(0.0526)	(0.1493)
Interest ceiling deregulation			0.0086	0.0022			0.0266	-0.0095
			(0.0085)	(0.0089)			(0.0453)	(0.0819)
Bank-insurance separation deregulation			0.0178***	0.0108*			0.2294**	0.1649**
			(0.0061)	(0.0056)			(0.0936)	(0.0727)
IPO share of market capitalization		0.0054**		0.0064***		0.0942***		0.0672**
		(0.0026)		(0.0023)		(0.0269)		(0.0263)
Corporate Default Rate		0.0027*		0.0034**		0.0814***		0.0720***
		(0.0016)		(0.0014)		(0.0213)		(0.0176)
Foreign Assets/GDP		-0.0146		-0.0769***		-0.6536***		-0.2378
		(0.0188)		(0.0254)		(0.1782)		(0.3073)
Top Marginal Tax Rate		-0.0009		-0.0062		0.0190		-0.1378
		(0.0098)		(0.0096)		(0.1009)		(0.1092)
Observations	96	84	96	84	98	84	98	84
R-squared	0.856	0.847	0.868	0.867	0.685	0.787	0.742	0.808
Sample	1910-2005	1920-2003	1910-2005	1920-2003	1909-2006	1920-2003	1910-2005	1920-2003

Notes. Standard errors in parentheses; Newey-West standard errors with 5 lags of autocorrelation. *** p<0.01, ** p<0.05, * p<0.1. The regulation variables are coded such that higher values indicate more deregulation (looser regulation).

Table A4: Market Capitalization and Executive Compensation

	Dependent Variable: Log Executive Compensation											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
		All Firms			Nonfarm Private Sector			Finance			All Firms	
Log market capitalization	0.31***	0.33***	0.38***	0.32***	0.34***	0.41***	0.26***	0.27***	0.26***	0.33***	0.35***	0.42***
	(0.005)	(0.005)	(0.009)	(0.006)	(0.005)	(0.010)	(0.018)	(0.017)	(0.029)	(0.006)	(0.006)	(0.009)
Log market capitalization of reference firm	0.60***	0.57***	0.50***	0.63***	0.60***	0.50***	0.25***	0.24***	0.26***	0.55***	0.52***	0.46***
	(0.017)	(0.017)	(0.020)	(0.020)	(0.020)	(0.023)	(0.040)	(0.039)	(0.047)	(0.018)	(0.018)	(0.021)
Log market capitalization X finance indicator interaction										-0.11***	-0.11***	-0.22***
										(0.015)	(0.014)	(0.026)
Finance indicator										0.41***		
										(0.142)		
Constant	0.36***			0.04			3.46***			0.66***		
	(0.138)			(0.159)			(0.343)			(0.143)		
Firm random effects	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Industry fixed effects	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes	No
Firm fixed effects	No	No	Yes	No	No	Yes	No	No	Yes	No	No	Yes
Observations	15,375	15,375	15,375	13,864	13,864	13,864	1,511	1,511	1,511	15,375	15,375	15,375
Number of firms	2,403	2,403	2,403	2,175	2,175	2,175	228	228	228	2,403	2,403	2,403
Overall R-squared	0.417	0.483	0.410	0.405	0.476	0.396	0.382	0.438	0.382	0.406	0.476	0.213

Notes. Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1. In columns 1-9 the reference is the median firm within each sample in each year. In columns 10-12 the reference firm is the median within finance or nonfarm private sector (excluding finance) for firms in each respective sector. Executive compensation data are from Execucomp and market capitalization is from Compustat in 1992-2010. We start with the 1000 top firms by CEO compensation before taxes and drop firms in agriculture and firms with no market capitalization information. We convert nominal values into 2010 prices using the Consumer Price Index series from the Bureau of Labor Statistics.

Table A5: Regressions used for estimating wage profiles

	Dependent variable: log hourly wage					
Sample:	Men, experience<=5			Men, experience<=29		
	1971-1980	1981-1990	1991-2005	1971-1980	1981-1990	1991-2005
White indicator	0.0586***	0.0482***	-0.00307	0.125***	0.111***	0.0842***
	(0.00745)	(0.00797)	(0.00627)	(0.00346)	(0.00329)	(0.00249)
Married indicator	0.0898***	0.0754***	0.0766***	0.121***	0.121***	0.146***
	(0.00421)	(0.00504)	(0.00511)	(0.00252)	(0.00232)	(0.00194)
Urban indicator	0.00985**	0.0120**	-0.00284	0.00484**	-0.00295	-0.0156***
	(0.00453)	(0.00526)	(0.00482)	(0.00229)	(0.00239)	(0.00203)
Education = 12 years	0.119***	0.118***	0.113***	0.205***	0.242***	0.267***
	(0.00621)	(0.00739)	(0.00729)	(0.00276)	(0.00324)	(0.00309)
Education = 13 to 15 years	0.222***	0.271***	0.254***	0.314***	0.390***	0.438***
	(0.00679)	(0.00801)	(0.00727)	(0.00325)	(0.00358)	(0.00318)
Education = 16 years	0.386***	0.550***	0.564***	0.509***	0.620***	0.730***
	(0.00720)	(0.00818)	(0.00794)	(0.00339)	(0.00367)	(0.00335)
Education > 16 years	0.516***	0.714***	0.777***	0.621***	0.790***	0.978***
	(0.00992)	(0.0113)	(0.0110)	(0.00433)	(0.00436)	(0.00388)
Probability of unemployment	0.704***	1.347***	-0.154	0.765***	0.972***	-0.193***
	(0.0684)	(0.0826)	(0.120)	(0.0324)	(0.0341)	(0.0454)
Experience^2	-0.00716***	-0.00119	0.00228**	-0.00108***	-0.000913***	-0.000911***
	(0.000862)	(0.00100)	(0.000898)	(1.61e-05)	(1.70e-05)	(1.43e-05)
Experience	0.102***	0.0669***	0.0367***	0.0516***	0.0491***	0.0460***
	(0.00498)	(0.00589)	(0.00512)	(0.000498)	(0.000534)	(0.000462)
Experience X Finance indicator	-0.00689	0.000410	0.0245***	0.00562***	0.00397***	0.00477***
	(0.00745)	(0.00776)	(0.00715)	(0.000649)	(0.000665)	(0.000553)
Finance indicator	0.0514**	0.107***	0.0864***	0.0216**	0.105***	0.156***
	(0.0262)	(0.0277)	(0.0254)	(0.0104)	(0.0107)	(0.00959)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Observations	34,650	32,454	42,189	186,652	218,034	359,764
R-squared	0.223	0.274	0.266	0.287	0.315	0.336

Notes. Data: March CPS. Top coded wages are multiplied by 1.75. All workers are full time full year male employees, who earned at least 80% of the federal minimum hourly wage. The probability of unemployment next year conditional on employment in the current year is estimated using the Matched CPS. All regressions include year dummies. A constant was included, but not reported.

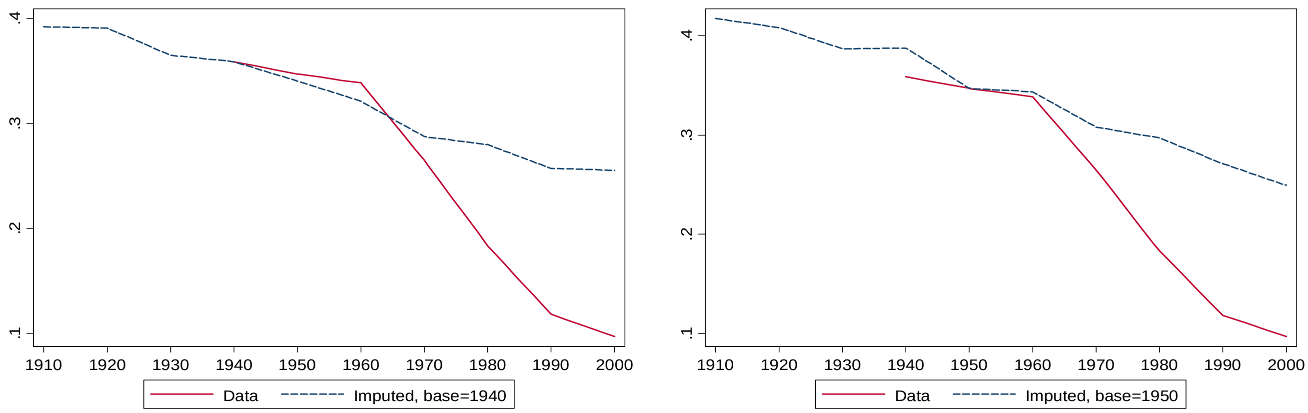
Table A6: Predicted wage volatility in finance versus non-finance

Sample:	Nonfarm Private Sector			Finance		
	1971-1980	1981-1990	1991-2005	1971-1980	1981-1990	1991-2005
Experience						
0	0.356	0.411	0.464	0.293	0.450	0.607
1	0.358	0.390	0.441	0.322	0.467	0.501
2	0.370	0.391	0.438	0.388	0.358	0.486
3	0.374	0.410	0.428	0.359	0.420	0.478
4	0.369	0.411	0.438	0.337	0.421	0.450
5	0.375	0.424	0.442	0.368	0.421	0.505
6	0.388	0.426	0.448	0.348	0.407	0.511
7	0.396	0.435	0.465	0.392	0.427	0.510
8	0.399	0.439	0.467	0.398	0.449	0.508
9	0.402	0.441	0.470	0.403	0.468	0.534
10	0.402	0.443	0.469	0.452	0.448	0.569
11	0.408	0.449	0.481	0.467	0.483	0.559
12	0.410	0.446	0.488	0.409	0.517	0.558
13	0.418	0.455	0.490	0.443	0.521	0.627
14	0.421	0.458	0.490	0.478	0.552	0.600
15	0.417	0.459	0.502	0.423	0.531	0.609
16	0.433	0.465	0.511	0.510	0.582	0.604
17	0.432	0.475	0.508	0.522	0.539	0.588
18	0.431	0.470	0.509	0.488	0.540	0.625
19	0.439	0.476	0.515	0.507	0.520	0.675
20	0.442	0.482	0.519	0.501	0.537	0.599
21	0.436	0.487	0.520	0.558	0.553	0.631
22	0.441	0.465	0.526	0.528	0.546	0.647
23	0.449	0.487	0.525	0.510	0.606	0.653
24	0.445	0.488	0.531	0.563	0.617	0.655
25	0.445	0.492	0.537	0.512	0.576	0.645
26	0.445	0.493	0.536	0.526	0.660	0.648
27	0.453	0.497	0.533	0.559	0.618	0.643
28	0.451	0.496	0.528	0.549	0.565	0.629
29	0.445	0.495	0.525	0.527	0.554	0.667
Average	41.49%	45.51%	49.15%	45.46%	51.18%	58.40%
Finance minus nonfarm				3.97%	5.66%	9.26%

Notes. The table reports standard deviations of residuals from the regressions reported in Table A5, for the experience<=29 years sample. The standard deviations are calculated separately for finance and the rest of the nonfarm private sector. Since the left hand side variable is log hourly wages, then the standard deviations can be interpreted in percent terms. Data: March CPS.

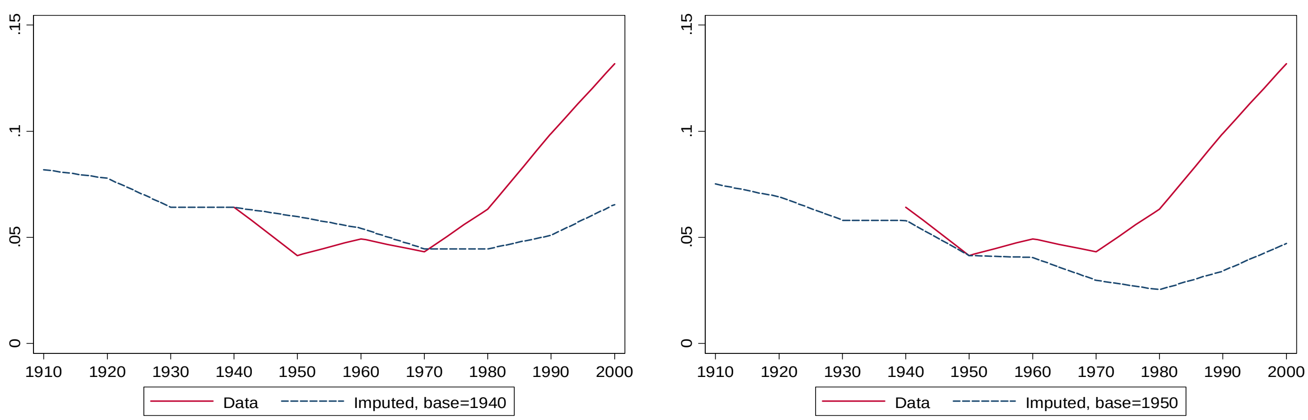
Figure A1: Imputations of Relative Educational Shares in Finance

A. Workers with 12 or More Years of Education



B. Workers with Strictly More than 12 Years of Education

C. Workers with 16 or More Years of Education



Notes. The solid lines are the education share in finance minus the education share in the nonfarm private sector; the sources are the 1940-2000 U.S. censuses. The dashed lines are imputed relative shares. The imputation uses educational shares within occupations in the base year (1940 or 1950). We use the 1950 consistent occupational classification available in IPUMS. The series in Figure I uses the imputed shares of strictly more than 12 years of education based on 1950 data in 1910-1930, and the actual data from 1940 and on.

Figure A2: Finance Relative Top Wage and Top Human Capital

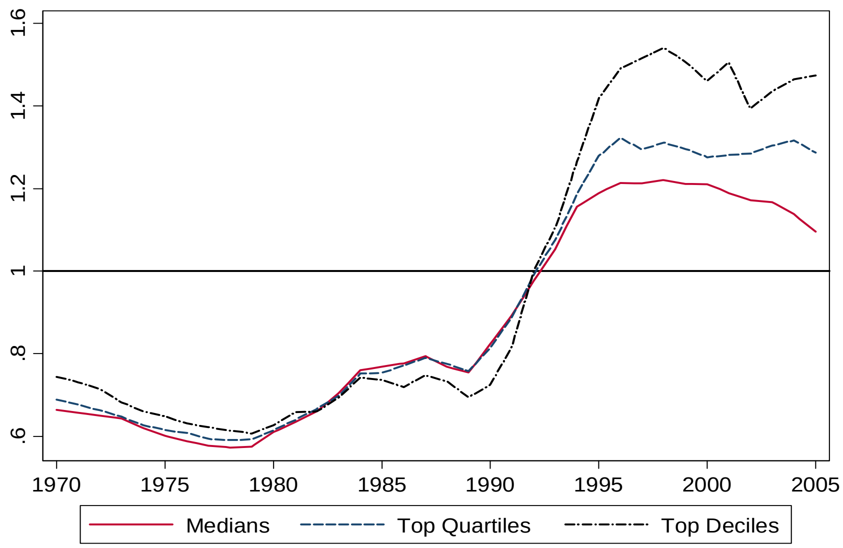
A. Using Top Quartiles

B. Using Top 50%

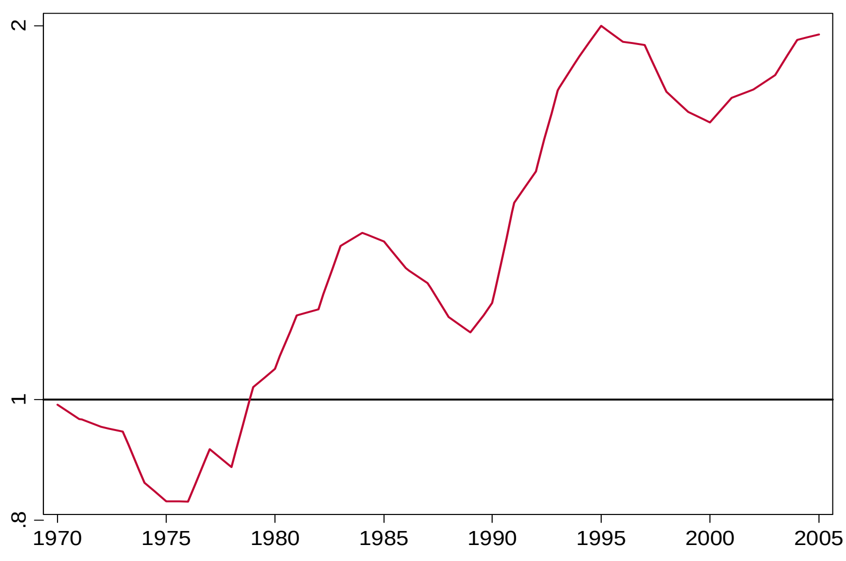
Notes. Top quartile or 50% human capital is the share of workers with wage above the top quartile or median wage in the nonfarm private sector including finance, respectively. Relative human capital is the difference in top quartile or 50% human capital shares between finance and the nonfarm private sector excluding finance, respectively. The top quartile or 50% average wage is the average wage of workers in the top quartile or 50% within some sector, respectively. It is computed by using the average wage below the top quartile or median from the U.S. censuses and the overall average wage from Figure I. The relative top quartile or 50% wage is the ratio of the top quartile or 50% average wage in finance to nonfarm private sector top quartile or 50% average wage, respectively. See text for complete details on the calculations.

Figure A3: Relative Executive Compensation in Forbes-Execucomp

A. Entire Sample



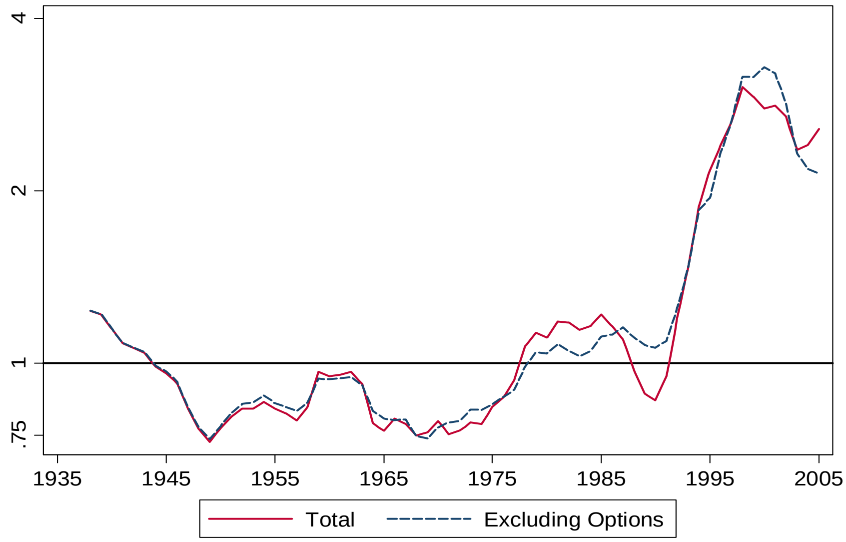
B. Top 50 Executives



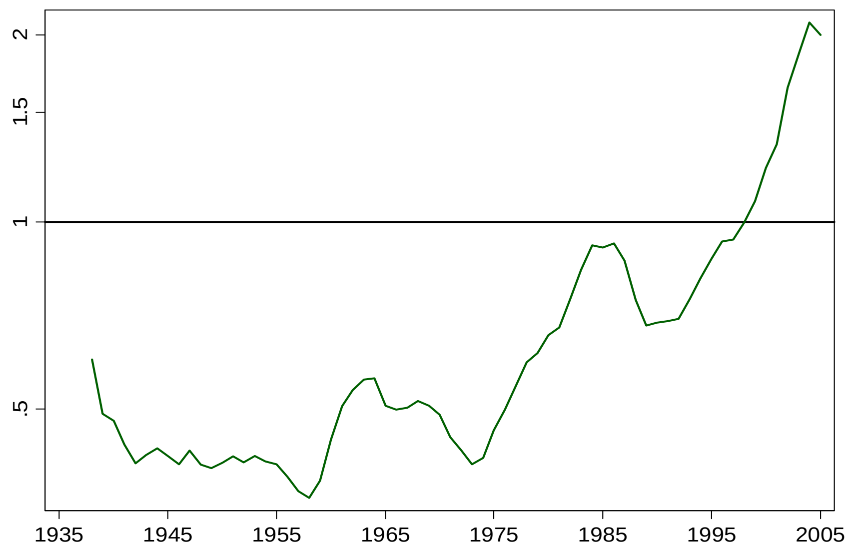
Notes. The source for all series is the Forbes-Execucomp sample, as described in the text. Panel A reports the ratio of median, top quartile and top decile CEO compensation in finance to the respective percentiles in the nonfarm private sector. Panel A uses the entire Forbes-Execucomp sample. The series in Panel B is ratio of the median compensation for the 5 highest paid CEOs in finance relative to the median for the 45 highest paid CEOs in the nonfarm private sector. Note that this is not the same as the 50 largest firms by market value, so the sampling methodology differs from that of Frydman and Saks (2007). In Panel B the vertical axis is log scale.

Figure A4: Relative Executive Compensation and Market Values

A. Median Executive Compensation in Finance Relative to Private Sector: Top 50 Firms

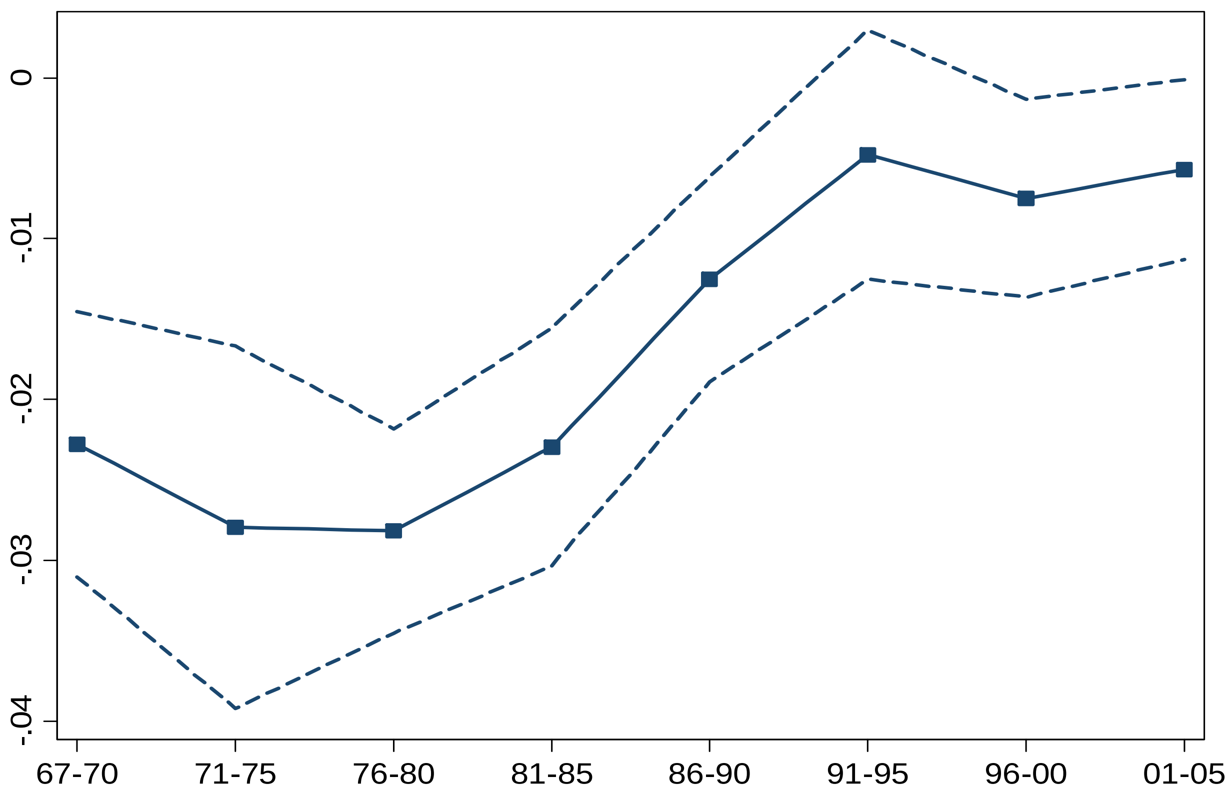


B. Median Market Value in Finance Relative to Private Sector: Top 50 Firms



Notes. Panel A presents median executive compensation in finance relative to median executive compensation in the rest of the nonfarm private sector. It is the same as Figure IV. Data are smoothed using a 5-year moving average. The vertical axis is log scale. The sample is the top three executives in each of 50 of the largest publicly traded firms that operated in the U.S. in 1936-2005 and reported executive compensation for at least 20 years, obtained from Frydman and Saks (2007). See their data appendix for complete documentation. None of these 50 firms are in agriculture, and 7 are in finance: CIT Group 1938-1976, Aetna 1964-2005, AIG 1970-2005, Citicorp (Citigroup) 1971-1997, Chase (J.P. Morgan Chase) 1972-2005, American Express 1977-2005, Cigna 1982-2005. The solid line take into account total executive compensation, including the value of options at the time they were granted estimated by the Black-Scholes formula. The dashed line excludes the value of options. The vertical axis is log scale. Panel B presents median market value of financial firms relative to median market value in the rest of the private sector. Data are smoothed using a 5-year moving average. The vertical axis is log scale. The sample includes 50 firms which reported market capitalization for at least for 20 years, of which the number of firms in finance follow the sampling in Frydman and Saks (one firm in 1938-1963, two in 1964-1969, three in 1970, four in 1971, five in 1972-1981 and 1998-2005, and six in 1982-1997). All firms are the largest in their subsector (finance or nonfarm private sector). Market values are taken from the Center for Research in Securities Prices database (CRSP).

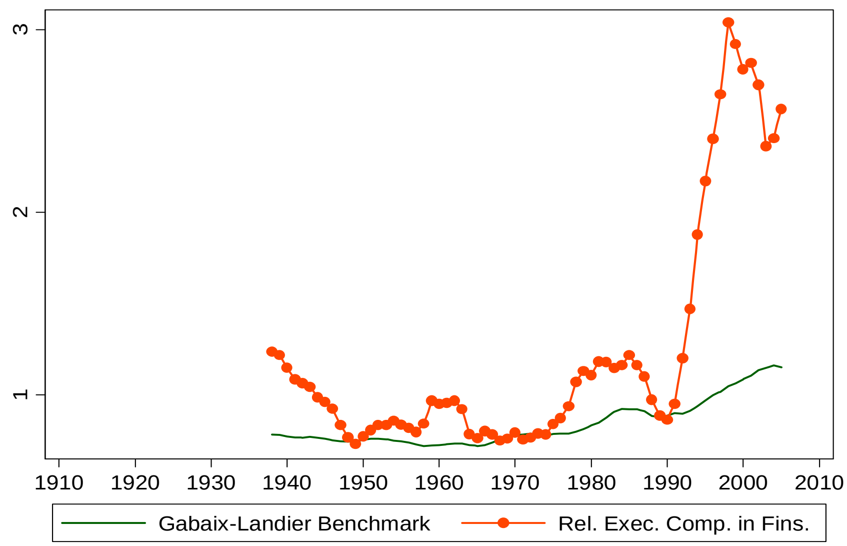
Figure A5: Relative Unemployment Risk in Finance



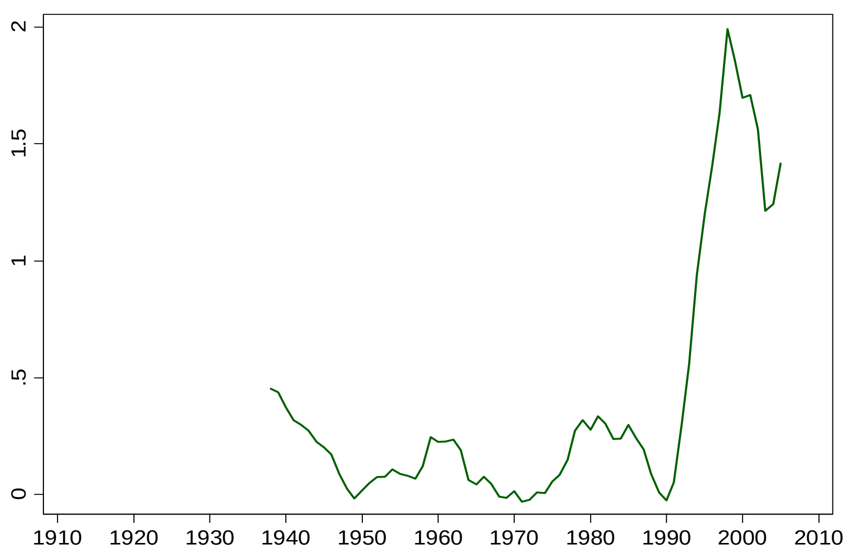
Notes. Coefficients and 95% confidence intervals of Finance dummy in logit regression of transition from Employment to Unemployment. Controls include current log hourly wage, race, sex, marital status, urban residence, potential experience and its square and education controls. Data: Matched CPS.

Figure A6: Relative Executive Compensation and Segregated Markets Benchmark

A. Relative Executive Compensation and Benchmark, Assuming Segregated Labor Markets



B. Excess Relative Executive Compensation



Notes. Panel A reports the relative executive (total) compensation in the financial industry (fins.), based on data from Frydman and Saks (2007), from Figure IV. The benchmark relative executive compensation series is constructed by applying the Gabaix and Landier (2008) methodology, under the assumption that labor markets for executives in finance and nonfarm private sector are segregated. It involves applying the coefficients from column 10 in Table A3 to firm market value data from the Center for Research in Securities Prices (CRSP). We adjust the benchmark so that it equals the relative compensation in 1966-1975 on average. The sample of firms is the same as in Frydman and Saks (2007). Data are smoothed using a 5-year moving average. Panel B reports the difference between the two series in Panel A: the excess relative executive compensation.