Internet Appendix to
SORTING IN THE U.S. CORPORATE EXECUTIVE LABOR MARKET

Egor Matveyev
MIT Sloan School of Management
Massachusetts Institute of Technology

August 2017

Contents

This appendix provides additional details on the construction of my sample, supplementary results, and technical details on estimation procedures.

A Sample Construction and Processing

A.1 Compensation Data

The main source of executive compensation data is Standard & Poor's ExecuComp database. The updated data files used in this version of the paper were downloaded in June 2016.

Firms report executive compensation, typically in Definitive Proxy Statements Schedule 14A, according to Item 402 of Regulation S-K. Firms must present a Summary Compensation Table (SCT), which includes three years of compensation data for each Named Executive Officer (NEO).[1] The SCT contains the following data:

1. **Salary and Bonus**
   - Salary and bonus must reflect the total salary and bonus amounts earned by each NEO during the respective year. Any amounts deferred, whether pursuant to a plan established under Section 401(k) of the Internal Revenue Code, or otherwise, should be included in the appropriate column for the fiscal year in which earned.
2. **Stock Awards**
   - Stock awards must disclose the dollar value of stock-related awards (other than "option awards" that are disclosed separately) that derive their value from the registrant's equity securities. "Stock" is defined as instruments such as common stock, restricted stock, restricted stock units, phantom stock, phantom stock units, common stock equivalent units or any other similar instruments that do not have option-like features. [Item 402(a)(6)(i) of Regulation S-K]
3. **Option Awards**
   - Option awards must disclose the dollar value of option-related awards. The term "option" is defined as instruments such as stock options, stock appreciation rights (SARs) and similar instruments that have option-like features [Item 402(a)(6)(i) of Regulation S-K].

[1] NEOs include: (1) the registrant's principal executive officer (CEO) and principal financial officer (CFO) or any individual(s) acting in those capacities at any time during the year; (2) the registrant's three highest paid executive officers, other than the CEO and CFO, who were serving as executive officers at the end of the last completed fiscal year and (3) up to two additional executive officers who would have met the requirements in (2), but for the fact that the individual was not serving as an executive officer of the registrant at the end of the last completed fiscal year.

4. **Non-equity Incentive Plan Compensation**

   - Must report the dollar value of all other amounts earned by the NEO during the fiscal year pursuant to non-equity incentive plans of the registrant. Awards are reported only if the incentive plan's relevant performance measure is not based on the price of the registrant's equity securities or the award may not be settled by issuance of the registrant's equity securities. Performance-based compensation under an incentive plan that is not indexed to or settled in the registrant's stock will be disclosed under this item in the year when the plan's specified performance criteria are satisfied and the compensation is earned (whether or not payment is actually made in that year).

5. **Change in Pension Value and Nonqualified Deferred Compensation Earnings**

   - This component must disclose the sum of: (1) the aggregate increase in actuarial present value of each NEO's accumulated benefit under all defined benefit and actuarial plans (including supplemental plans) during the year and (2) any earnings on nonqualified deferred compensation. Footnote identification and quantification of the separate amounts attributable to (1) and (2) above are required.

6. **All Other Compensation**

   - All other compensation to an NEO during the fiscal year that has not been reported in another column of the SCT must be disclosed in the All Other Compensation column. The only exception is if perquisites and other personal benefits for the NEO aggregated to an amount less than $10,000 during the year, in which case disclosure of such amount may be omitted. Each item of compensation included in this column for the last fiscal year that exceeds $10,000 for each NEO must be separately identified and quantified in a footnote to the table. Conversely, any individual item of compensation that is less than $10,000 for each NEO will be included in the All Other Compensation column but does not require separate identification and quantification in a footnote to the table.

7. **Total**

   - The total compensation reported in SCT represents the sum of the above.

The main annual compensation variable I employ in this paper is the total compensation reported in SCT as discussed above. This value corresponds the total dollar value earned by the NEO in a given year, and therefore represents rents to executive skill.

Since 2006 when the new set of SEC rules took effect, Execucomp introduced variable TOTAL_SEC, which captures exactly the total compensation reported in SCT. This data item is my main compensation variable after 2006. Before 2006, Execucomp was reporting the total compensation granted in a given year in item TDC1. According to Execucomp's definition, this item includes: "Salary, Bonus, Other Annual, Total Value of Restricted Stock Granted, Total Value of Stock Options Granted (using Black-Scholes), Long-Term Incentive Payouts, and All Other Total". All together, these items would represent the current definition of TOTAL_SEC, although separately items comprising TDC1 may be logged under different component in the SCT. Therefore, from 1992 until 2005 (inclusive), my main compensation variable is TDC1.

Although the data quality is generally considered very good, there are multiple inconsistencies and typos that need to be cleaned out. First, there are 39 entries with negative values of TOTAL_SEC. I have manually checked these numbers through EDGAR to verify that these are typos.[2]

Next, in 110 cases the value of TOTAL_SEC is reported zero. Out of 110 cases, 18 cases report positive TDC1. For all of these cases, the actual value of total compensation reported in the SCT corresponds to TDC1. Therefore, these 18 observations are replaced with TDC1. After 2006, there are 1,609 observations for which TOTAL_SEC is missing but TDC1 is positive. I replace all 1,609 observations with the value of TDC1.

In addition to TDC1, Execucomp reports a variable called TDC2. The difference between TDC1 and TDC2 is that TDC2 replaces the Black-Scholes value of options granted in a given year with the value of options exercised during that year. To be specific, the definition of TDC2 is as follows: "Salary + Bonus + Other Annual + Restricted Stock Grants + LTIP Payouts + All Other + Value of Options Exercised". In the cases where both TOTAL_SEC and TDC1 are missing, I use TDC2 to fill out missing observations. Post-2006, there are 18 such cases where I have used the TDC2 value.

Prior to 2006, I use TDC1 as my main compensation variable. In 24,156 cases where the TDC1 value is missing, TDC2 is reported. In all such cases, I use the value of TDC2 as NEO wage.

Importantly, I do not apply the procedure described above to zero values of TDC1. In particular, cases where TDC1 is reported as 0, and the value of TDC2 is reported as a positive number, I do not use TDC2 instead of TDC1. Most of these cases represent the NEO who has left the firm in the current fiscal year or the year before. Since there is no identifying information in compensation values that equal to zero, I delete such observations. There were 225 such cases. At the end of the processing, there were 49 observations left with missing values; all such observations were deleted.

A.2 Executives employed by more than one firm in a given fiscal year

I start my analysis with 254,011 observations. Out of these, 250,453 (98.6%) have a single annual executive-firm record. This means that these managers have been employed by only one firm in a

[2] In principle, a negative total value in SCT is possible if, for example, the NEO has not received any pay in a given year and the value of pension went down.

Name	Firm Name	Title	Year	TDC1
Raymond V. Gilmartin	BECTON DICKINSON & CO	pres. & CEO	1992	825.00
Raymond V. Gilmartin	BECTON DICKINSON & CO	chmn., pres. & CEO	1993	1,773.79
Raymond V. Gilmartin	BECTON DICKINSON & CO	chmn., pres. & CEO	1994	5,993.00
Raymond V. Gilmartin	MERCK & CO	chmn., pres. & CEO	1994	7,589.47
Raymond V. Gilmartin	MERCK & CO	chmn., pres. & CEO	1995	3,991.28
Raymond V. Gilmartin	MERCK & CO	chmn., pres. & CEO	1996	4,927.89
Raymond V. Gilmartin	MERCK & CO	chmn., pres. & CEO	1997	7,567.20
Raymond V. Gilmartin	MERCK & CO	chmn., pres. & CEO	1998	9,019.65
Raymond V. Gilmartin	MERCK & CO	chmn., pres. & CEO	1999	17,916.69
Raymond V. Gilmartin	MERCK & CO	chairman, president & CEO	2000	12,478.39
Raymond V. Gilmartin	MERCK & CO	chairman, president & CEO	2001	17,737.73
Raymond V. Gilmartin	MERCK & CO	chairman, president & CEO	2002	12,888.87

Table 1: Example of a Transition Year

given fiscal year. For the remaining observations, executives are on record in more than one firm. In 1,755 instances executives are employed by two firms in a given year, and in 16 cases there are three executive-firm records in a given year.

An example is given in Table 1. It shows that in year 1994, Raymond V. Gilmartin is said to be the CEO of both Becton Dickinson and Merck. This record is suggestive of Gilmartin moving from Becton to lead Merck. This is confirmed by DowJones Newswire on 9 June 1994: "Merck & Co. (MRK) named Raymond V. Gilmartin president, chairman and chief executive. The company said Gilmartin will replace P. Roy Vagelos as president and chief executive, effective June 16 and as chairman in November, when Vagelos will formally retire. Gilmartin is currently chairman, president and chief executive of Becton Dickinson & Co. (BDX)."

I call a year in which an executive switches jobs a 'Transition Year'. I use the following definition to identify transitions in the data. I search for executives employed by two firms in a given year. Suppose executive A is employed by firms X and Y in year T. Year T is a transition year for executive A if the following holds:

1. A is employed only by X it year $T-1$ and only by Y in year $T+1$,
2. A is employed only by X it year $T+1$ and only by Y in year $T-1$.

I handle these transition years by keeping only one observation where the executive earned most of her income. In the example above, if the compensation paid to A by X is greater than that paid to her by Y, I keep only record of A employed by X in year T. If instead of X it was firm Y that paid A highest compensation, I keep Y record and delete X. In Gilmartin's case, I keep only Merck observation in year 1994. Becton observation from 1994 is deleted.

Name	Firm Name	Title	Year	TDC1
Laurence Alan Tisch	CBS INC	chmn., pres. & CEO	1992	2,228.33
Laurence Alan Tisch	CNA FINANCIAL CORP	CEO	1992	21.50
Laurence Alan Tisch	LOEWS CORP	chmn. & co-CEO	1992	397.83
Laurence Alan Tisch	CBS INC	chmn., pres. & CEO	1993	2,740.80
Laurence Alan Tisch	CNA FINANCIAL CORP	CEO	1993	21.50
Laurence Alan Tisch	LOEWS CORP	chmn. & co-CEO	1993	613.35
Laurence Alan Tisch	CBS INC	chmn., pres. & CEO	1994	3,270.14
Laurence Alan Tisch	CNA FINANCIAL CORP	CEO-CNA Financial Corp.	1994	26.00
Laurence Alan Tisch	LOEWS CORP	co-chmn. & co-CEO	1994	649.59
Laurence Alan Tisch	CNA FINANCIAL CORP	CEO-CNA Financial Corp.	1995	26.00
Laurence Alan Tisch	LOEWS CORP	co-chmn. & co-CEO	1995	921.83
Laurence Alan Tisch	CNA FINANCIAL CORP	CEO-CNA Financial Corp.	1996	26.00
Laurence Alan Tisch	LOEWS CORP	co-chmn. & co-CEO	1996	1,874.45
Laurence Alan Tisch	CNA FINANCIAL CORP	CEO-CNA Financial Corp.	1997	26.00
Laurence Alan Tisch	LOEWS CORP	co-chmn. & co-CEO	1997	2,175.26

Table 2: Laurence Tisch's Employment Record: Original

Unlike transition years, there are patterns in the data where an executive is on record at two or more firms for more than one year. One example is Laurence Tisch. Table 2 shows that Tisch is employed at three firms: CBS, CNA, and Loews in years 1992-1994. He is employed in Loews and CNA during 1995-1997.

Most of such cases are unique in the data and I handle them manually. It turns out that Tisch is de facto one of the founders of Loews Corporation. CNA is an insurance firm and owned by Loews. In 1986 he acquired a 25% stake in CBS, and became its CEO, while retaining posts of co-CEO of Loews and CEO of CNA. He spent most of his time on CBS's business, which he left in 1995 after Westinghouse Electric bought CBS. After that, he concentrated his effort on managing Loews. He formally resigned in 2000. In this case it is reasonable to assume that most of the rents to his talent accrue from CBS in years 1992-1994, and from Loews in years 1995-1997. Therefore I collapse this record as shown in Table 3.

At the conclusion of the procedure, I am left with 252,211 compensation records.

A.3 Construction of tenure variable

Execucomp has two variables that capture when the executive has joined the firm and when she has become (if at all) the CEO. Variable 'JOINED_CO' captures the date when the executive has joined the firm. Variable 'BECAMECEO' gives the date when the executive became the CEO. I employ the following algorithm to construct tenure variable used in this paper.

Name	Firm Name	Title	Year	TDC1
Laurence Alan Tisch	CBS INC	chmn., pres. & CEO	1992	2,228.33
Laurence Alan Tisch	CBS INC	chmn., pres. & CEO	1993	2,740.80
Laurence Alan Tisch	CBS INC	chmn., pres. & CEO	1994	3,270.14
Laurence Alan Tisch	LOEWS CORP	co-chmn. & co-CEO	1995	921.83
Laurence Alan Tisch	LOEWS CORP	co-chmn. & co-CEO	1996	1,874.45
Laurence Alan Tisch	LOEWS CORP	co-chmn. & co-CEO	1997	2,175.26

Table 3: Laurence Tisch's Employment Record: Processed

There are four possible options with respect to individual's tenure with a firm. First, a person can be identified as a CEO in every year of her employment with the firm. I refer to these individuals as "Always CEOs". Second, a person can be identified as any executive other than the CEO in each year of her employment. I refer to them as "Never CEOs". Third, a person can start her job at a firm as an executive, and during her tenure can be promoted to the CEO position. These group of executives I call "internally promoted CEOs". Fourth, and final option, is a CEO who retires (or get demoted) but remains with the firm in a role other than the CEO. I call this group "retired CEOs". I discuss how I handle these cases in turn.

1. "Always CEOs": These individuals are identified as CEOs in "CEOANN" variable.
   (a) If 'BECAMECEO' is available, I use this date to identify the CEO's tenure.
   (b) If 'BECAMECEO' is not available, but 'JOINED_CO' is available, I take the midpoint since the time the CEO has joined the firm. The reason for this choice is follows. I have manually checked 30 cases in which 'JOINED_CO' was available and 'BECAMECEO' was not. In 8 out of 30 cases, 'JOINED_CO' date was actually the date when the executive has become the CEO. (Most of these cases were external hires.) In the remaining cases, the executive was internally promoted to the CEO position, but Execucomp did not code 'BECAMECEO' variable. In these 22 cases, the distribution of when the executive has become the CEO relative to joining the firm, was closely approximating the uniform distribution. Based on this evidence, I have chosen the midpoint as the expected value of when the executive was promoted to the CEO position. One important note: since in some cases, as I discussed above, the 'JOINED_CO' is the actual date when the executive has become the CEO, this procedure biases the true tenure down.[3]
   (c) If neither 'BECAMECEO' nor 'JOINED_CO' are available, the CEO's tenure cannot be identified.

[3]The main results in the paper are not affected and not sensitive to dropping these observations from the analysis.

2. "Never CEOs": This group of people is identified as non-CEOs in "CEOANN" variable.
   (a) For these individuals, I define tenure as the length of the overall employment with the firm, regardless of the position that the person was employed in.[4]
   (b) If 'JOINED_CO' is available, I use this date to identify the executive's tenure.
   (c) If 'JOINED_CO' is not available, tenure cannot be identified.
3. Internally promoted CEOs: Their employment records start by them identified as non-CEOs in "CEOANN", and during the tenure "CEOANN" identifier switches to "CEO"
   (a) To be consistent with above definitions, I break down tenure definition for these executives into two:
      i. First, for the years until the executive becomes a CEO, tenure is defined as the overall length of executive's employment with the firm.
         - Construction of the variable it identical to the "Never CEOs" class.
      ii. Second, in the first year they become a CEO, I reset they tenure to zero, and start counting tenure with the firm as the CEO.
         - Construction of the variable it identical to the "Always CEOs" class.
   (b) Conceptually, this definition determines tenure for CEOs starting from the point of their employment in the CEO role. For all other executives, tenure is determined since the time they have originally joined the firm.
4. Retired CEOs: Their employment records indicate them as CEOs in "CEOANN", which at some point switches to non-CEO.
   (a) For the years these executives are CEOs, the tenure variable is defined identical to the "Always CEOs" class.
   (b) For the years after they retire (or get demoted), I continue counting their tenure as if they were continuing as CEO.
   (c) Economically, I assume that tenure will affect pay for this class of executives similar to the way it affected pay when they were CEOs.

A.4 Construction of age variable

Executive age in ExecuComp is given by two variables. One is called 'AGE' and is meant to capture executive's age at the time of filing of the proxy statement or annual report. The second

[4] For example, if Apple hires a junior designer, and this person makes it to become the VP of Design, her tenure would be counted as starting from the date of her initial employment at Apple.

one is called ‘PAGE’ and is meant to capture the present age of the executive. Perhaps surprisingly, there are lots of missing observations and data inconsistencies.

I employ the following algorithm to construct the age variable used in this paper:

- For each executive
  - Locate her full employment record,
    - Locate the first firm executive worked
      1. Examine the record by locating the most recent year T
      2. If ‘AGE’ variable is available, use it to extrapolate the age on all the years in this employment record,
      3. If the ‘AGE’ variable at T is not available
         (a) Locate year $T-1$. If ‘AGE’ variable is available, use it to extrapolate the age on all the years in this employment record,
         (b) Repeat until the very first year in the employment record until the ‘AGE’ is located.
      4. If ‘AGE’ is not available for any year in the employment record, locate ‘PAGE’.
         (a) If ‘PAGE’ is available, I extrapolate executive’s present age on each year of her employment record. I assume ‘PAGE’ is given as of 2015.
         (b) If ‘PAGE’ is not available, continue onto the next employment record
    - Continue to the next firm; repeat steps 1-4 above,
  - If there are firm records in which ‘AGE’ and ‘PAGE’ are not available, but one of these variables is available for another firm where this executive has worked, extrapolate age.
  - Check for internal consistency. If executive’s age appears to be inconsistent, use the latest available year for which the age is available, and extrapolate that age on all firms in the employment record,
- Continue until all executives in the dataset are processed.

A.5 Missing CEOs

There are X out of 44,042 firm-year observations for which Execucomp does not identify firms’ CEOs. For these observations, I designate a highest paying individual in a firm as a CEO.

I randomly selected 50 out of X instances and manually checked whether the designated individual was the CEO. I use Definitive Proxy DEF 14A statements from SEC’s EDGAR system. In all 50 cases, the designated individual was the firm’s CEO.

A.6 Matching to BoardEx based on names

The ExecuComp and BoardEx databases each have their own set of ID numbers referring to unique individuals. To cross reference the datasets, individual names in each database have to be converted into the same format, and then matched.

The following is the list of conversion steps.

- For the ExecuComp database:
  - The full name is contained in the 'exec_fullname' variable
  - Split the full name string into multiple columns by the spaces
  - Manually move or combine name components to ensure they are correctly lined up
- For the Boardex database
  - The full name is contained in the 'Full Name' variable
  - Split the full name by the spaces
  - Sort the split data by the number of columns
  - Combine the individual's nickname and middle name/initial where applicable
  - Extract the middle initial from the combined nickname and middle name column. The logic is as follows,
    - For each cell in the middle name column check for a "("
      - If there is a "(", take the first character after ") " [note the space after the closed parenthesis]
      - If there is not a "(", take the first character
    - Enter the chosen character in a new "Middle Initial" column
  - Manually move or combine name components to ensure they are in the correct column (i.e. if there is only a first name and last name, move the first name from column 1 to column 2 and the last name from column 2 to column 5)

The matching of the stripped names is done as follows:

- For each of the 45,985 names in the ExecuComp database
  - Check the list of last names from the Boardex database. If there is a match, continue, if not, move to the next name in the ExecuComp list,
    - Check if the ExecuComp first name matches the Boardex first name for the identified match. If there is a match, continue, if not, move to the next name in the ExecuComp list,

- Check if the first character in the ExecuComp middle name matches the Boardex middle initial for the identified match. If there is a match, continue, if not, move to the next name in the ExecuComp list,
- If a match is identified, copy over the relevant Boardex ID number into the ExecuComp list,
- If not, continue onto the next ExecuComp ID.

This code matched 13,053 records with 695 duplicate matches (due to identical names) among them. These duplicates were manually corrected in one of two ways:

1. Middle name, title, or suffix information distinguished the two individuals and the correct Boardex ID numbers were assigned (101 records), or

2. The individuals were indistinguishable and the Boardex ID number was deleted from both records (594 records).

ExecuComp-BoardEx ID link is available upon request.

A.7 Education Data

Median college entrance year is 1971.
25th percentile is 1964
75th percentile is 1977
I use years 1971 and 1980.
1967 falls into 34th-36th percentile of the distribution. 1980 falls into 83rd-85th percentiles.
1977-and above (top 25th percentile, use 1980)
Table IA.4
Table **??**.

The definition of competitiveness groups is given in Table IA.5.

The remaining colleges in the U.S. are considered Non-Competitive or Non-Selective. These colleges admit many of the students with SAT scores in the lower tail of the distribution. I do not include specialized colleges, such as art schools, music schools, U.S. military academies, into the analysis. All such colleges were assigned a code of '1'.

B Identification of Fixed Effects

B.1 Preliminaries

In the AKM model

$$\log w_{it} = \alpha_i + \psi_{J(i,t)} + X_{it}'\beta_1 + Y_{J(i,t)}'\beta_2 + \varepsilon_{it}.$$

person effects α and firm effects ψ may not be separately identified. To see this, imagine the structure of the labor market where executives never move between firms. If all executives work for just one firm during their careers, separate identification of manager and firm effects is not possible. The identification requirement, therefore, requires the existence of job mobility in the data.

Consider the following example. For simplicity, assume that compensation is given net of observables X_{it} and $Y_{J(i,t)}$ affecting managers' pay. The market consists of two firms, A and B, and three executives, 1, 2, and 3. Executive 2 works for both firms A and B.[5]

Firm	Executive	Compensation	Firm effect	Person effect
A	1	\$100	0	0
A	2	\$150	0	50
B	2	\$170	20	50
B	3	\$150	20	30

Table B.4: Example 1: Labor market structure

The difference between the compensation Executive 2 receives in firm A compared to firm B implies that firm B pays \$20 more compared to firm A. This identifies the relative difference between firm fixed effects, i.e., $\psi_B - \psi_A = \$20$. There is no additional information that would allow us to identify firm A and firm B fixed effects separately, therefore, we have to normalize one of the effects. For example, if we set $\psi_A = 0$, then $\psi_B = 20$.

Person effects can be identified similarly. In particular, notice that Executive 2's fixed effect should be \$50 more than 1's: $\delta_2 - \delta_1 = 50$. Also, $\delta_2 - \delta_3 = 20$. Similar to identification of firm effects above, there is no additional information that can allow us to identify all three person effects separately. Normalizing $\delta_1 = 0$, we get $\delta_2 = 50$ and $\delta_3 = 30$.

The above example shows that in this labor market, both firm and executive fixed effects are identifiable up to a constant. This example features a group of firms that I will call **connected**. I discuss the algorithm that I use to construct connected groups in my dataset in Section *B*.3. A formal definition of a connected group is given below.

Definition A.1: Consider a group of firms that consists of J firms. This group is called **connected** if for any firm j_k in the group there exists at least one firm j_l such that firms j_k and j_l have at least

[5]In practice, executives rarely work for more than one firm at the same time. Exceptions mostly include founders. For example, Elon Musk as of 2016 is the CEO of Tesla and the CEO of SpaceX. In this example I abstract from such situations. More generally, I discuss how I handle such situations in the Data section of the paper.

one executive that worked for both of these firms.

Give the above definition, I can formally show that it is enough to normalize one firm and one person fixed effect, and the rest of the effects will be identifiable relative to this normalization.

Proposition A.1: In a connected group of firms, the sets of both firm and person fixed effects are identifiable up to a constant. In a group of J firms and I individuals, $J-1$ firm effects and $I-1$ person effects are identifiable. In total, $J+I-2$ firm and person effects are identifiable.

Proof: Start with the simplest connected group that has two firms, 1 and 2, such as the one shown in the example above. This group is connected according to Definition A.1 because Executive 2 works for both firms A and B. As Table $B.4$ shows, we can normalize one fixed effect, say ψ_1, and ψ_2 will be identified from the executive who works for both firms. Note that by Definition A.1 of a connected group such an executive exists. Now we add Firm 3 to this set of connected firms. Since Firm 3 is connected to firms 1 and 2, it means that at least one of firm's 3 executive works in either firm 1 or 2. Without loss of generality, suppose that she works for firm 2. The difference between her compensation in firm 2 relative to Firm 3 can be used to identify the effect of Firm 3. The same argument can be applied for any firm j that we add to this connected group of firms. This proves that in a connected group of J firms, we can separately identify $J-1$ firm effects. ■

B.2 Identifying Information in Group Means

Consider the following example with two separately connected groups of firms. The first group consists of firms A and B, as in example $B.4$. The second group consists of firms C and D. Note that none of the executives who works in firms A and B, works for either firm C or D. Therefore groups 1 and 2 are not connected.

If we set $\psi_C = 0$ and $\delta_4 = 0$, then we can get $\psi_D = 30$, and $\delta_5 = -50$ and $\delta_6 = -90$, similarly to the example in $B.4$.
Note that the pay in the second group is on average higher than the pay in the first group. It means that firms in the second group are high paying firms or executives at these firms have higher ability. Economically, it is interesting to understand why firms in certain connected groups choose to pay their executives more. I discuss the economics behind this choice in Section X of the paper. Econometrically, I show below that group means contain additional identifying information.

Let M_1 denote the mean salary in the first group of firms and M_2 in the second. In the example in Table $B.5$, $M_1 = 142.5$ and $M_2 = 267.5$. I set the average firm fixed effect in each group equal to

				Effects		Normalized effects	
Group	Firm	Executive	Compensation	Firm	Person	Firm	Person
1	A	1	$100	0	0	132.5	−32.5
1	A	2	$150	0	50	132.5	17.5
1	B	2	$170	20	50	152.5	17.5
1	B	3	$150	20	30	152.5	−2.5
2	C	4	$300	0	0	252.5	47.5
2	C	5	$250	0	−50	252.5	−2.5
2	D	5	$280	30	−50	282.5	−2.5
2	D	6	$240	30	−90	282.5	−42.5

Table B.5: Example 2: Labor market structure

mean pay in the group. In particular, I set

$$M_1 = \frac{\psi_A + \psi_B}{2}. \tag{1}$$

Equation (1) together with $\psi_B - \psi_A = 20$ exactly identifies both effects:

$$2M_1 = \psi_A + \psi_B = 2\psi_A + 20$$
$$\psi_A = M_1 - 10 = 132.5.$$

This gives $\psi_B = 152.5$. The normalization procedure for the second group is identical. It gives $\psi_C = 252.5$ and $\psi_C = 282.5$.

The above procedure allows me to use group means as additional data points to strictly identify all firm effects in the full sample. Importantly, notice that this procedure accomplishes two things. First, it makes group differences in average pay to be inherited in average group effects. This becomes relevant in the analysis of firm effects. Second, person effects are free of any group-specific firm unobservable components of pay. This information is useful to keep in mind when we economically interpret person effects.

B.2.1 No constant in the main equation

I estimate the AKM model by dropping the constant. Doing this makes it possible to exactly identify every person effect in each group. Therefore, I completely eliminate the need for mechanical normalization of effects.

Recall that in Example 1 I had to set $\delta_1 = 0$. If there is no constant in the model, the sum of firm and person effect for each pair must equal to the person's compensation. This helps with the identification of δ_1. If we take (A,1) pair, then

$$100 = \psi_A + \delta_1$$
$$\delta_1 = 100 - 132.5 = -32.5.$$

The rest of the effects in the group are identified relative to δ_1 as before. Note that since the average firm effects within groups equal to the mean group pay, it implies that the average (and the sum) of within-group person effects equals zero. I formally prove this proposition below.

Proposition A.2: Within-group sum of person fixed effect equals to zero.

Proof: The proof is straightforward. ■

Proposition A.3: In a sample that consists of G groups, I executives, and J firms, the total of $I+J-G$ fixed effects are identifiable.

Proof: The proof is straightforward. ■

B.3 Construction of connected groups

A *connected* group of managers and firms contains all firms where each manager worked during her career and all managers that firms ever employed. In contrast, a group of executives and firms is *not connected* to another group if no firm in the first group ever employed an executive in the second group and any executive in the first group was never employed by a firm in the second group. A simple example is shown in Figure 1.

The following algorithm constructs G mutually-exclusive groups of connected observations from I managers in J firms observed over the sample period. Start with an arbitrary manager. Include all firms in which he was ever employed. Next, add all executives who currently work or who have ever worked for these firms. Continue adding all additional firms for which any of these managers has ever worked (or currently works) and all additional managers in any of those firms until no more managers or firms can be added to the current group. Repeat for the next group and continue until no more observations left. At the conclusion of the algorithm, the persons and firms in the sample have been divided into G groups. The number of managers in each group is I_g. The number of firms in each group is J_g. Within each group g, $I_g - 1 + J_g - 1$ person and firm effects are identified. Overall in all G groups, exactly $I+J-G$ effects are estimable. This assertion can be formally proved.

?

B.4 Normalization across groups

Given the above identification procedure, there is no additional normalization across groups that is required. The grand sum of estimated manager effects across years in the sample equals to zero.

If person and firm fixed effects are estimated with unknown normalization process, for example routine REGHDFE in Stata, I provide the following simple process that helps to normalize the estimated effects.

Start with firm effects. Denote by ψ_j the set of firm effects estimated with "canned" procedure, and denote $\widetilde{\psi}_j$ the set of normalized effects. Calculate the first firm effect from the following equation:

$$M = \sum_g \psi_j - T_g \psi_1 + T_g \tilde{\psi}_1$$
$$\tilde{\psi}_1 = \frac{1}{T_g}\left(M - \sum_g \psi_j + T_g \psi_1 \right).$$

where M is the sum of wage residuals within group, not counting person and firm effects, i.e., the sum of $\log w_{it} - X_{it}'\hat{\beta}_1 - Y_{J(i,t)}'\hat{\beta}_2$, and T_g is the number of observations in the group.

Once $\tilde{\psi}_1$ is known, all other $\tilde{\psi}_j$ can be computed from the following equation:

$$\tilde{\psi}_j - \tilde{\psi}_1 = \psi_j - \psi_1.$$

Similar normalization procedure can be applied to estmated person effects. Denote by α_i the set of person effects estimated with "canned" procedure, and denote $\tilde{\alpha}_i$ the set of normalized effects that we would like to use. Calculate the first firm effect from the following equation:

$$0 = \sum_g \alpha_i - T_g \alpha_1 + T_g \widetilde{\alpha}_1$$
$$\widetilde{\alpha}_1 = \frac{1}{T_g} \sum_g \alpha_i + \alpha_1.$$

Each subsequent person effect is computed as:

$$\widetilde{\alpha}_i - \widetilde{\alpha}_1 = \alpha_i - \alpha_1.$$

B.5 Accounting for co-worker effects in the statistical model

As I discuss . Here I show the convergence properties of the procedure.

C Bias Correction

Covariance of estimated worker-coworker effects is positively biased.

Simulate on the given structure of the market.

Cite **?** . Now

Cite **?**

C.1 Inference

My statistical inference is based on the simulated distribution of the test statistics. This subsection provides important details on the procedure.

The bias in the

C.1.1 Data generating process

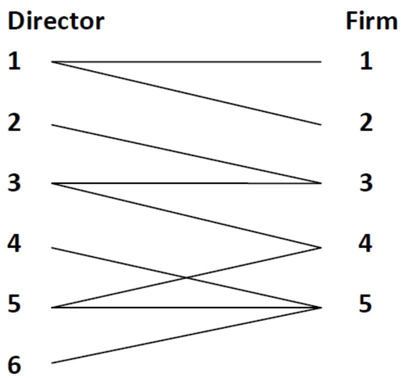
I simulate firm effects with the mean and

D Computation

No inversion. Use Congugate algorithm.

Following Card I estimate the model using Congugate. It converges but does not give exact estimates. I apply the following normalization procedure.

Figure 1: Example of the grouping procedure



Director	Firm	Group
1	1	1
1	2	1
2	3	2
3	3	2
3	4	2
4	5	2
5	4	2
5	5	2
6	5	2

Table IA.1: MINCER RETURNS TO EACH COLLEGE CATEGORY

The table shows results from OLS **?** regressions. The dependent variable is the natural logarithm of executive compensation. Variables R2 through R9 define college or university rank. Rank R1 (Non-Competitive) serves as a benchmark. Graduate degree and (Graduate degree)∗MBA are dummy variables defining graduate non-MBA and MBA degrees, respectively. Standard errors are clustered at individual person level.

	College Rank	Highest Rank
R2	-0.002	0.041
	(0.085)	(0.093)
R3	0.078	0.051
	(0.071)	(0.083)
R4	0.121	0.077
	(0.079)	(0.090)
R5	0.193***	0.169**
	(0.072)	(0.084)
R6	0.150**	0.101
	(0.089)	(0.098)
R7	0.232***	0.227***
	(0.075)	(0.086)
R8	0.269**	0.207**
	(0.112)	(0.094)
R9	0.353***	0.354***
	(0.077)	(0.083)
Graduate degree	-0.009	
	(0.035)	
(Graduate degree)*MBA	0.069***	
	(0.026)	
Age	0.188***	0.187***
	(0.014)	(0.013)
Age^2	-0.002***	-0.002***
	(0.000)	(0.000)
Constant	1.040***	1.078***
	(0.364)	(0.359)
Observations	70,634	73,626
R^2	0.142	0.143

Table IA.2: NUMBER OF EXECUTIVES PER FIRM BY YEAR

The table shows the distribution of the number of executives per firm in the sample. The second and third columns show the number of unique firms and executives each year. Columns (4) and (5) shows the minimum and the maximum executives per firm, and columns (6) and (7) report means and medians, respectively.

(1) Year	(2) Firms	(3) Executives	(4) Min	(5) Max	(6) Average	(7) Median
1992	1,571	8,022	1	11	5.11	5
1993	1,682	9,757	1	12	5.80	6
1994	1,749	10,609	1	12	6.07	6
1995	1,848	11,044	1	15	5.98	6
1996	1,974	11,552	1	15	5.85	6
1997	2,034	11,940	1	15	5.87	6
1998	2,069	12,501	1	14	6.04	6
1999	1,952	12,078	1	12	6.19	6
2000	1,846	11,455	1	13	6.21	6
2001	1,846	11,314	1	13	6.13	6
2002	1,884	11,487	1	14	6.10	6
2003	1,933	11,746	1	13	6.08	6
2004	1,878	10,859	1	12	5.78	6
2005	1,761	9,349	1	12	5.31	5
2006	1,924	10,877	1	11	5.65	5
2007	2,176	12,320	1	12	5.66	5
2008	2,094	11,793	1	12	5.63	5
2009	2,049	11,227	1	13	5.48	5
2010	2,013	10,839	1	13	5.38	5
2011	1,964	10,580	1	11	5.39	5
2012	1,913	10,321	1	12	5.40	5
2013	1,859	9,992	1	13	5.37	5
2014	1,818	9,474	1	11	5.21	5
2015	205	1,075	1	9	5.24	5

Table IA.3: EMPLOYMENT STRUCTURE OF THE U.S. EXECUTIVE LABOR MARKET

	Group 1	Group 2	Other groups	Total	Percent in Group 1
Number of groups	1	1	1,151	1,153	
Firms	2,203	7	1.1	3,514	62.7%
Executives	31,502	83	11.0	44,299	71.1%
Total observations	177,672	550	63.9	252,211	70.4%
Estimable effects	33,704	89		46,660	72.2%

Notes: Groups are ordered from largest to the smallest by the number of individuals in every group. For example, "Group 1" has the largest number of executives identified by procedure described in the text.

Table IA.4: Distribution of imputed college entrance years

Birth year	Occurrence	Percent	Birth year	Occurrence	Percent
1924	4	0	1965	6,944	3.4
1925	7	0	1966	7,424	3.64
1926	7	0	1967	7,011	3.44
1928	23	0.01	1968	7,159	3.51
1929	5	0	1969	7,484	3.67
1930	4	0	1970	8,097	3.97
1931	40	0.02	1971	7,843	3.84
1932	33	0.02	1972	8,068	3.95
1933	47	0.02	1973	8,311	4.07
1934	96	0.05	1974	8,738	4.28
1935	34	0.02	1975	8,159	4
1936	54	0.03	1976	8,253	4.04
1937	149	0.07	1977	7,694	3.77
1938	36	0.02	1978	7,831	3.84
1939	108	0.05	1979	6,875	3.37
1940	105	0.05	1980	5,732	2.81
1941	254	0.12	1981	5,246	2.57
1942	167	0.08	1982	4,949	2.42
1943	188	0.09	1983	3,980	1.95
1944	321	0.16	1984	3,262	1.6
1945	420	0.21	1985	2,678	1.31
1946	502	0.25	1986	2,390	1.17
1947	550	0.27	1987	1,804	0.88
1948	783	0.38	1988	1,436	0.7
1949	1,111	0.54	1989	1,137	0.56
1950	885	0.43	1990	817	0.4
1951	1,058	0.52	1991	505	0.25
1952	1,598	0.78	1992	446	0.22
1953	1,747	0.86	1993	261	0.13
1954	1,876	0.92	1994	159	0.08
1955	2,205	1.08	1995	128	0.06
1956	2,673	1.31	1996	104	0.05
1957	2,716	1.33	1997	48	0.02
1958	3,099	1.52	1998	26	0.01
1959	4,046	1.98	1999	23	0.01
1960	4,672	2.29	2000	5	0
1961	4,544	2.23	2001	3	0
1962	5,022	2.46	2002	4	0
1963	5,140	2.52	2004	1	0
1964	6,719	3.29	2005	2	0

Table IA.5: Barron's definition of competitiveness and assignment rules

Rank	Category	1971	1980
9	Most Competitive	Colleges in this group require high school rank in the 10% to 20% and grade averages of A to B+. Median freshman test scores at these colleges are between 675 and 800 on the SAT and above 28 on the ACT.	Colleges in this group require high school rank in the 10% to 20% and grade averages of A to B+. Median freshman test scores at these colleges are between 625 and 800 on the SAT and above 28 on the ACT.
8	Highly Competitive Plus	Colleges in this group look for students with grade averages of B+ and B and accept most of their students from the top 20% to 30% of the high school class. Median freshman scores close to 675 on the SAT or close to 28 on the ACT and colleges that accept fewer than one-quarter of their applicants.	Colleges in this group look for students with grade averages of B+ and B and accept most of their students from the top 20% to 30% of the high school class. These are colleges with median freshman scores close to 625 on the SAT or close to 28 on the ACT and colleges that accept fewer than one-quarter of their applicants.
7	Highly Competitive	Colleges in this group look for students with grade averages of B+ and B and accept most of their students from the top 20% to 30% of the high school class. Median freshman test scores at these colleges range from 600 to 675 on the SAT and from 26 to 28 on the ACT.	Colleges in this group look for students with grade averages of B+ and B and accept most of their students from the top 20% to 30% of the high school class. Median freshman test scores at these colleges range from 575 to 625 on the SAT and from 26 to 28 on the ACT.
6	Very Competitive Plus	The colleges in this category admit students whose averages are no less than B- and who rank in the top 30% to 50% of their graduating class. They accept fewer than one-third of their applicants and colleges with median freshman test scores near the top of the "Very Competitive" range.	The colleges in this category admit students whose averages are no less than B- and who rank in the top 30% to 50% of their graduating class. These colleges accept fewer than one-third of their applicants and colleges with median freshman test scores near the top of the "Very Competitive" range.
5	Very Competitive	The colleges in this category admit students whose averages are no less than B- and who rank in the top 30% to 50% of their graduating class. They report median freshman test scores in the 550 to 600 range on the SAT and between 23 and 26 on the ACT.	The colleges in this category admit students whose averages are no less than B- and who rank in the top 30% to 50% of their graduating class. They report median freshman test scores in the 525 to 575 range on the SAT and between 23 and 26 on the ACT
4	Competitive Plus	Colleges in this category report median freshman SAT scores well over 500 and median freshman ACT scores well above 22, those that admit fewer than one-half of their applicants, or those with admission requirements that make them more competitive than the other institutions in the Competitive category.	Colleges in this category report median freshman SAT scores well over 500 and median freshman ACT scores well above 22, those that admit fewer than one-half of their applicants, or those with admission requirements that make them more competitive than the other institutions in the Competitive category.
3	Competitive	The category is a wide one, covering colleges with median freshman test scores from the upper 400's (above 450) to about 550 on the SAT and 21 to 23 on the ACT. Many of these colleges require that students have high school averages of B- or better, although other colleges state a minimum of C+ or C. Generally, these colleges prefer students in the top two-thirds of the graduating class.	The category is a wide one, covering colleges with median freshman test scores from about 425 to 525 on the SAT and 20 to 23 on the ACT. Many of these colleges require that students have high school averages of B- or better, although other colleges state a minimum of C+ or C. Generally, these colleges prefer students in the top two-thirds of the graduating class.
2	Less Competitive	Included in this category are colleges with median freshman test scores below 450 on the SAT and below 21 on the ACT; many colleges that require entrance examinations but do not report median scores; and colleges that admit students with C averages who rank in the top 75% of the graduating class.	Included in this category are colleges with median freshman test scores below 425 on the SAT and below 20 on the ACT; many colleges that require entrance examinations but do not report median scores; and colleges that admit students with C averages who rank in the top 75% of the graduating class.