Initial Industry Choices & Long-Term Growth in Earnings

Stephen L. Ross* Patralekha Ukil[#]

Abstract

This paper examines the relationship between the initial industry choice made by an individual, and the exposure from that choice on long term earnings growth, while conditioning on a large number of variables that give us information on the individual's demographic characteristics, family background and measures of ability. The analysis uses data on male workers from three distinct cohorts of the National Longitudinal Survey of Youth and historical industry data from the Current Employment Statistics of the Bureau of Labor Statistics. The results suggest that there is a significant, positive and persistent relationship between making a "favorable" industry choice and growth in earnings. When examined in greater detail, the results also indicate that the effects are concentrated among occupation types which involve more routine and manual tasks, and also among those that involve fewer abstract tasks. Interestingly, this relationship does not depend on the persistence or "length of stay" within the initial industry.

Keywords: Industry, Industry Characteristics, Job Choice, Wage Differential, Skill Biased, Earnings Inequality

JEL Classification: J21, J22, J24, J31, J62

*Department of Economics, University of Connecticut, 365 Fairfield Way, Unit 1063, Storrs CT 06269-1063

[#]Department of Economics, San Jose State University, One Washington Square, San Jose CA 95192-0114

Section I: Introduction

Numerous individual and labor market characteristics may determine the long-term earnings trajectory of an individual, and it is of considerable interest to economists and policy makers to understand how different factors influence long-term earnings. The initial industry choice of an individual may exert a significant influence by creating a positive or a negative "shock" on the long-term earnings profile. It is quite reasonable to think that individuals with similar occupations and similar skills who enter into "favorable" industries as opposed to "unfavorable" industries early in their careers, may end up with very different earnings trajectories over their lifetimes.

The initial industry choice of a young worker might be a critical determinant of later-life labor market outcomes. This choice affects the firm-specific or industry-specific human capital accumulated by the worker, which in turn could determine their relative bargaining position within the labor market. It could lead to significant persistence of earnings or wages across jobs, and could also act as a signal of skills to an employer, or even determine the worker's risk of lay-off or job displacement. Thus, one's initial industry choice could set in place a trajectory of job conditions which would impact the long-term income profiles of a worker.

This paper analyses the relationship between the industry chosen by an individual at entry into the labor market and long-term growth in earnings, using data from three different cohorts of the National Longitudinal Survey of Youth: the NLSY Young Men Cohort, the NLSY79 and the NLSY97. This is an unexplored research question and the existing literature lacks adequate empirical evidence about this relationship, especially in the context of the United States. While examining the relationship between early industry choices and long-term earnings growth, this analysis also tries to evaluate how much this relationship depends on the initial (or pre-market) endowment that each individual brings into the labor market: years of schooling, measures of ability, cognitive, non-cognitive and social skills, parental or family background and regional characteristics, etc. We control for worker attributes which are observed either prior to or during the initial industry choice; each observed effectively 10 years or 20 years prior to the labor market outcomes being considered.

Labor market choices tend to be influenced by a large number of observable and unobservable characteristics, and it is important to understand whether the association between industry choice and long-term earnings growth gets depleted *when* we control for an individual's initial or pre-labor market endowment.

The results indicate that, on average, earnings growth is significantly positively associated with making a 'favorable' industry choice. Conditional upon a large variety of individual characteristics, measures of skills and ability and family background characteristics, the results show that earnings growth is significantly higher when workers enter into industries

which expanded or grew considerably over time. The coefficient of interest remains stable and statistically significant even after controlling for demographic characteristics, education, measures of ability, family background and regional characteristics and after including different fixed effects. Furthermore, this significant and positive relationship often persists for 20 years from that initial industry choice, and the results are robust to several different specifications. Heterogeneity analysis shows that the association between making a favorable initial industry choice and earnings growth *does not depend* on persistence or the period of time a worker spends within the initial industry – results do not differ for those who left the initial industry relatively early, compared to those who persisted within the initial industry for at least five years.

We conduct further analysis in which we look at the effects on categories of workers based on their occupational task content. We find that although the differences between the groups are not significant, the effects are concentrated among occupation types which involve more routine and manual tasks, and also among those that involve fewer abstract tasks. We also analyze results based on the offshorability factor of the occupational tasks, and find the effects to be more concentrated among occupations that can be more easily offshored. Though the differences in the subgroups in the heterogeneity analyses are not statistically significant, the patterns are suggestive and are consistent with the existing literature on the polarization of wages and rising earnings inequality within the United States.

This paper proceeds as follows: Section II reviews the current literature related to the research question; Section III explains the empirical methodology; Section IV describes the data used and discusses descriptive statistics; Section V reports the main results of the paper, including several checks on the robustness of the results, and also provides the results from several heterogeneity analyses. Section VI concludes the paper.

Section II: Related literature

The research in this paper relates to two different but very associated strands of literature – the first one being on the factors which influence long-term labor market outcomes, and the second one being on the issue of the increasing earnings inequality.

Earlier research on efficiency wages, such as Dickens & Katz (1987) and Krueger & Summers (1988), found substantial dispersion in wages across industries which were often measured by the standard deviation in industry wage differentials. The analyses found very little difference between union and non-union workers, and these large industry effects were observed even after controlling for individual characteristics.

Gardecki & Neumark (1998) examine the impact of initial labor market mobility or stability on later-life labor market outcomes, to understand whether faster transitions to stable employment would lead to better labor market outcomes later on. The authors find modest evidence of a positive relationship between early labor market stability and better later-life labor market outcomes. Specifically, they do not find any strong relationship between early

labor market stability and better labor market outcomes for men. For women, on the other hand, they find weak evidence of positive effects of early labor market stability.

In a related context, many authors have examined the relationship between the state of the labor market at the time of entry and subsequent labor market outcomes. Oyer (2006) finds that the macroeconomic conditions prevalent at the time of graduation from an Economics Ph.D. program affect both the short-term and long-term career prospects of economists. Similarly, Oyer (2008) shows that stock market shocks have persistent effects on the careers of MBA graduates, and these shocks influence whether or not the graduates go to Wall Street after graduation. Kahn (2010) finds large and negative effects on wages of graduating in a bad economy which persist for many years. Oreopoulous et al. (2012) study the impact on earnings of graduating from college during a recession using employer-employee matched data from Canada, and find negative effects which last up to 10 years. Schwandt & von Wachter (2017) also find an adverse persistent effect of entering into the labor market during a recession, and that these negative effects are larger for those without a college degree and also for non-white workers.

Several other papers have studied both the short-term and long-term effects of job displacement or job loss on later life labor market outcomes. For the U.S., using administrative data on workers and firms, Jacobson et al (1993) estimate both the magnitude and the pattern of displaced workers' earnings losses and find that high-tenure workers who separate from firms face a long-term loss in earnings of approximately 25% per year. Similarly, Couch & Placzek (2010) use administrative data from the state of Connecticut for the period 1993 to 2004, to study the effect of job displacement on the earnings losses of workers. They find 32-33% earnings losses for workers who face mass layoffs during a period of moderate labor market conditions. Burgess et al. (2003) study the impact of unemployment early in an individual's career on the employment prospects later on in life, using data from the United Kingdom. They find a persistent negative effect of early unemployment for unskilled individuals, while finding a small but positive impact for more skilled individuals. For Germany, von Wachter & Bender (2006) use administrative data on young German workers to examine long run impact of an early career job loss. Their results show that earnings losses for the young displaced workers are initially 15% but slowly reduce to zero within a period of five years. They also find that persistent earnings losses are suffered by those young workers who leave very large establishments.

A slightly different but related literature is the one on the increasing earnings inequality in the United States. Most of the earlier studies on this topic utilized public-use micro data samples from the Current Population Survey. Katz & Autor (1999) provide a very comprehensive summary of the literature that analyses the increase in the earnings inequality in the U.S., and document that earnings inequality increased substantially since the 1970s and mainly during the 1980s. Kopczuk et al. (2010) use data from the Social Security Administration to analyse earnings inequality and labor market mobility from 1937 to 2004, and find that the inequality over this period looks U-shaped – decreasing till 1953 and increasing after that. Subsequent literature has focused on the possible sources of this increasing earnings inequality and the reasons for the increasing divergence of the earnings

distribution: Acemoglu & Autor (2011) focus on skill-biased technological change as a source of earnings inequality, while Autor (2014) focuses on the increase in the returns to education and cognitive ability. Yet others have focused on the differential demand for and returns to occupational skills to explain the polarization of earnings in the U.S., such as increased computerization of the workplace and the reduced demand for routine skills (Autor et al., 2003; Autor & Dorn, 2013; Hanushek et al., 2015).

Several authors have also focused on the role of the employer or the firm in explaining this trend of growing earnings inequality. Majority of such studies use administrative data to explain this effect. Spletzer (2014) uses data from the Longitudinal Employer-Household Dynamics (LEHD) and shows that the increasing inequality in the U.S. from 1996 to 2012 was driven by a rising variation in the average wages across firms, rather than by a rising variation across workers within the same firm. The author also finds that quarterly earnings inequality has increased among those workers who have a larger attachment to the labor market. Card et al. (2013) use administrative data from Germany, focusing on full-year full-time male employees, and find that the increase in German wage inequality was partially due to the increase in the dispersion in the wage premiums paid by the firms. Barth et al. (2016) use several datasets such as the March CPS, the LEHD and the Longitudinal Business Database (LBD) to analyze the impact of establishments in earnings dispersion and find that the majority of the increase in earnings inequality in the 1970s-2010s resulted from the increased dispersion of earnings across the establishments where the individuals were working.

Mueller et al. (2017) examine the relationship between wage inequality and firm growth, and find that wage differentials between high-skill jobs and medium / low-skill jobs increase with the size of the firm. Abowd et al. (2018) analyze the role of the employer in earnings inequality, and find that there are significant gains from working at a top-paying firm for all types of skills. They find that working for a high-paying employer produces current benefits through higher earnings, which persist over the worker's lifetime via an increase in the likelihood of upward mobility in the workplace. Autor et al. (2017) analyze micro panel data since 1982 and document the phenomenon of the fall in the labor share and the rise of "superstar firms". They hypothesize that as globalization or technology progresses, the industries will become increasingly dominated by these superstar firms with large profits and a low share of labor in the value-added or sales. They find support for their predictions that industry sales will concentrate in a small number of firms and the industries with the highest concentration will show the largest declines in labor share, and the fall in the labor share will be largely driven by between-firm allocation rather than within firms.

Muller and Neubaeumer (2018) analyze how life-cycle unemployment of apprentices depend on the training firms' sizes. Using large administrative data from Germany they find that training at a larger firm leads to lower unemployment in the long run. In a very related context, Arellano-Bover (2019) uses social security data from Spain to investigate the long-term impact of starting a first job at a small firm versus a large firm. He finds that starting a career at a larger firm (higher number of employees) as opposed to a smaller one results in substantial positive impacts on long-term outcomes – the analysis finds an elasticity between lifetime income and first-employer size to be equal to 0.12. The author argues that larger firms

provide important channels of human capital formation to workers, which also lead to subsequent jobs at other large employers.

Thus, both in the context of the rising earnings inequality and factors which affect life-time earnings, the role of the initial industry is an important research question. However, the existing literature lacks a broad and overarching assessment of the relationship between the initial industry choice and later life labor market outcomes, especially in the context of the United States. *Our contribution to the literature is an empirical analysis of this critical early-career choice, which could determine the trajectory of a worker's earnings profile*: through its influence on the human capital accumulation and the relative bargaining position of workers within the labor market; which could lead to persistence of earnings across jobs or even act as a signal to employers determining workers' risk of job displacement, and could also be a significant factor contributing to the rising earnings inequality in the United States.

We carry out this analysis in the context of the U.S labor market, using longitudinal data from three separate cohorts of the National Longitudinal Survey of Youth. This dataset allows us to control for many of the typically unobservable individual characteristics that bias standard OLS estimates. We run several checks on our results to gauge their robustness to different specifications. We also attempt to understand how these results may be heterogeneous depending upon the persistence within the initial industry and the task content of the initial occupation that the worker enters into.

Section III: Empirical Methodology

The covariate of interest is a measure of making a "favorable" initial industry choice versus an "unfavorable" one. Thus, in this context, we first identify and differentiate a "favorable" initial industry from an "unfavorable" one. We use the growth rate of employment in the individual's initial industry (over 10 and 20 years from entry) as a ***proxy*** for the covariate of interest. The employment growth rate provides a measure of the overall demand for workers within that industry, and also a sense of the performance of the industry over time.

The initial industry is identified as the first industry which the individual reports being employed in, following full-time working hours. We identify the first NLSY survey year an individual is in full time employment, and in order to maximize the sample of analysis, we consider all individuals who have declared an industry within 7 years of starting full-time employment[1].

The estimating equation is as follows:

$$Y_i = \beta_0 + \beta_1 G_i + \beta_2 X_i + \gamma_t + \delta_r + \varepsilon_i$$

Y_i is the outcome of interest (growth in earnings) pertaining to individual i;

[1] The actual year of declaring the initial industry is referred to as the "**entry year**" in our analysis.

G_i refers to the employment growth in individual i's initial industry;

X_i relates to the different control variables which may explain the outcome of interest;

γ_t stands for *age-at-first-survey* by *entry-year* fixed effects[2] and δ_r stands for initial region of residence fixed effects. Here, β_1 is our coefficient of interest.

We estimate the above equation in a cross-sectional framework, where the dependent variable is the growth in (real) earnings[3] over 10 years and 20 years from the individual's entry into the initial industry. We perform a cross-sectional analysis, instead of a panel analysis, because the focus is on exploring the long-term influence on the growth in earnings, as opposed to exploring year-to-year variations in earnings.

Section IV: Data & Summary Statistics

IV.A: Data

The data for this paper comes from three different survey cohorts of the National Longitudinal Survey of Youth – the Young Men cohort (part of the NLSY Original Cohort project), the 1979 cohort and the 1997 cohort. The Young Men's cohort has longitudinal data on 5225 men who were between 14 and 24 years of age when they were first surveyed in 1966, with data available till 1981. The NLSY79 cohort includes respondents who were aged between 14 and 22 years during their first survey in 1979, with data available up to 2014. The NLSY97 cohort follows respondents who were aged 12-17 years when first interviewed in 1997, with data available up to 2016. Thus, all three surveys provide rich data on young workers in the U.S., from three separate decades, and they also allow us to follow these individuals over a significant portion of their working lives.

In this analysis, we restrict our attention to only male workers, who are not employed in the agricultural sector or the armed forces[4]. To obtain our sample of analysis: we first follow individuals over time to observe the year they declare full-time employment[5] and this year is considered the individual's entry year. The industry they declare being employed within in the entry year is considered their initial industry. Next, in order to maximize sample, we include all individuals who have declared an initial industry within 7 years of being in full-time employment. We proceed to remove any individuals with no earnings reported in the entry year, and also the years before and after the entry year. Then we proceed to remove all individuals with no valid observations of earnings and industry 10 years (and subsequently 20 years) after their entry year.

[2] Age-at-first-survey here refers to the age at which the respondent was first interviewed by NLSY. Year of entry FE is added to control for macroeconomic conditions during entry into industry.
[3] Dependent variable is calculated as growth in real earnings over 10 (& 20) years, with appropriate base years chosen for each NLSY cohort.
[4] We do this allow us to avoid any issues related to the different fertility and labor force participation decisions of women which generally influence female samples, and the different confounding factors which may impact the agricultural sector and the armed forces.
[5] Full-time employment implies working 30 or more hours per week

Each control variable included in our regression either pertains to the entry year computed for the individual or was observed before the individual entered the initial industry. For example, we consider the region of residence declared in the entry-year as the initial region of residence. Similarly, the family income declared at the entry year is considered the initial family income.

In order to calculate the outcome variables, we consider three-year moving average of earnings, and we remove those years in which earnings are either reported as zero or are unavailable. The earnings data are deflated by the appropriate indices and thus our outcomes correspond to growth in real earnings 10 years and 20 years from entry into the industry.

In order to calculate the proxy variable for the covariate of interest, i.e., the growth in the initial industry of each individual, we use industry employment data from the Current Employment Statistics (CES) program, which is an establishment payroll survey of the Bureau of Labor Statistics.

Variables indicating individual ability include total years of schooling, dummies for whether or not the individual has a high-school degree, a college degree, some college education, and also the individual ASVAB components of the Armed Forces Qualification Test (AFQT)[6]. In addition, for the 1979 cohort, we follow Deming (2017) and include relevant variables that indicate social skills and non-cognitive skills, such as the Rosenberg self-esteem score and the Rotter Locus of Control score (both standardized), self-reported sociability and participation in clubs[7].

IV.B: Summary statistics

Table 1 provides descriptive statistics for the sample of men who entered into industries which expanded (in terms of employment) over 10 years, and those which contracted over 10 years.

Table 1: Summary statistics by industry type

Variable	Individuals who entered expanding industries (Mean & S.E.)	Individuals who entered contracting industries (Mean & S.E.)	Difference
1966 (Young Men cohort)			
Black	0.23	0.23	0.0004
	(0.007)	(0.039)	(0.040)
Other race[8]	0.006	0.00	0.0065
	(0. 0015)	(0.00)	(0.007)
Age at initial survey	18.46	17.05	1.413***
	(0.058)	(0.247)	(0.296)

[6] We have ASVAB information only for the 1979 and the 1997 cohort. For the 1966 cohort we use the standardized form of a composite (IQ) score reported by the survey which combines the results from several aptitude and intelligence tests.
[7] Description of these variables are provided in the Appendix Table A.2.
[8] The 1966 survey does not classify "Hispanic" as a separate race, while the 1979 and the 1997 surveys do.

Total years of schooling	13.252 (0.050)	13.07 (0.264)	0.184 (0.257)
College graduate	0.26 (0.008)	0.23 (0.039)	0.0295 (0.0416)
High school graduate	0.34 (0.008)	0.34 (0.044)	-0.0076 (0.0447)
Some college education	0.23 (0.007)	0.28 (0.042)	-0.0423 (0.0401)
Father's years of education	9.20 (0.152)	10.67 (1.022)	-1.4627 (1.469)
Mother's years of education	10.17 (0.129)	10.17 (0.833)	0.0082 (1.311)
Initial family income ($)	3129.66 (56.7365)	3265.95 (270.4525)	-136.28 (288.148)
Initial region: Non-south	0.60 (0.009)	0.59 (0.045)	0.0153 (0.046)
Average real earnings during entry ($) (3-year avg.)[9]	3914.78 (54.565)	3902.05 (245.28)	12.736 (276.583)
Average real earnings 10 years from entry ($) (3-year avg.)	9246.99 (91.517)	9059.37 (454.55)	187.629 (465.492)
1979 cohort			
Black	0.28 (0.07)	0.22 (0.013)	0.060*** (0.016)
Hispanic	0.18 (0.006)	0.19 (0.013)	-0.008 (0.014)
Age at initial survey	17.51 (0.038)	18.17 (0.076)	-0.657*** (0.084)
Total years of schooling	13.29 (0.044)	12.55 (0.044)	0.738*** (0.096)
College graduate	0.23 (0.007)	0.16 (0.012)	0.073*** (0.0154)
High school graduate	0.86 (0.005)	0.81 (0.013)	0.052*** (0.0133)
Some college education	0.22 (0.007)	0.15 (0.012)	0.0685*** (0.015)
Father's years of education	11.219 (0.073)	10.262 (0.144)	0.957*** (0.161)
Mother's years of education	11.081 (0.057)	10.319 (0.114)	0.761*** (0.126)
Initial family income ($)	20984.78 (500.708)	21055.10 (585.232)	-70.321 (1005.446)
Initial region: North-East	0.19 (0.006)	0.16 (0.012)	0.030*** (0.014)
Average real earnings during entry ($) (3-year avg.)	4825.02 (76.01973)	5895.712 (156.5671)	-1070.692*** (168.6519)

[9] Average real earnings are calculated by taking a 3-year moving average of reported real earnings, both at the initial entry period and also the 10-year and 20-year period.

Average real earnings 10 years from entry ($) (3-year avg.)	26088.04 (386.057)	24704.6 (633.7135)	1383.434 (820.3886)

1997 cohort			
Black	0.20 (0.012)	0.15 (0.014)	0.0439*** (0.019)
Hispanic	0.19 (0.012)	0.19 (0.015)	-0.0003 (0.019)
Age at initial survey	15.03 (0.044)	15.39 (0.050)	-0.367*** (0.068)
Total years of schooling	14.02 (0.091)	13.57 (0.107)	0.458*** (0.142)
College graduate	0.35 (0.0152)	0.30 (0.0176)	0.0485*** (0.023)
High school graduate	0.45 (0.0159)	0.47 (0.019)	-0.028 (0.025)
Some college education	0.28 (0.014)	0.24 (0.016)	0.041 (0.022)
Father's years of education	13.078 (0.107)	12.460 (0.134)	0.618*** (0.171)
Mother's years of education	13.057 (0.090)	12.585 (0.106)	0.471*** (0.140)
Initial family income ($)	61868.81 (2393.998)	62883.55 (3216.738)	-1014.733 (3925.675)
Initial region: North-East	0.17 (0. 012)	0.17 (0. 014)	0.002 (0.018)
Average real earnings during entry ($) (3-year avg.)	9167.372 (244.2608)	11520.64 (355.5747)	-2353.263*** (416.7725)
Average real earnings 10 years from entry ($) (3-year avg.)	37311.49 (874.82)	40726.69 (1144.418)	-3415.204*** (1418.819)

*****Mean differs from the corresponding mean of the sample entering expanding industries at p<0.05 level.**

From the descriptive statistics, it is clear that men who entered into industries which expanded over 10 years have higher perceived levels of education and ability and also parental education. *Thus, it is motivating for us to question if the growth in earnings over 10 years can be mostly explained by their inherent ability and schooling, and whether the initial industry choice essentially is actually irrelevant in explaining earnings growth.*

Section V: Results

V.A: Main results - Explaining growth in earnings 10 years after entry year

Table 2 below shows the results of the main analysis, where the dependent variable in the growth in average (real) earnings over 10 years.

Table 2: Growth in earnings 10 years after industry entry

Growth in (log) average real earnings	**(1)**	**(2)**	**(3)**	**(4)**
1966 cohort:				
Growth in industry employment	0.816***	0.384***	0.227**	0.213*
	(0.166)	(0.123)	(0.106)	(0.105)
Observations	3,002	3,002	3,002	3,002
R-squared	0.285	0.389	0.460	0.461
1979 cohort:				
Growth in industry employment	0.815***	0.513***	0.509***	0.506***
	(0.144)	(0.108)	(0.102)	(0.102)
Observations	4,305	4,305	4,305	4,305
R-squared	0.105	0.223	0.237	0.241
1997 cohort:				
Growth in industry employment	0.969***	0.565***	0.518***	0.538***
	(0.241)	(0.126)	(0.124)	(0.126)
Observations	1,638	1,638	1,638	1,638
R-squared	0.147	0.272	0.279	0.282
Age-at-first-survey x Entry-Year FE	YES	YES	YES	YES
All demographic + education variables		YES	YES	YES
All family background variables			YES	YES
Initial Region FE				YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

NOTE: Each panel represents separate OLS regressions (standard errors in parentheses clustered by initial industry) for the NLSY cohort. Data is aggregated up to time-consistent two-digit industries.

We find a positive and significant relationship between the growth in an individual's earnings and the growth in the initial industry employment 10 years from the entry year, in all of the different specifications, and for all of the three NLSY cohorts.

Column (1) shows the estimate of the relationship without any controls, but with the age-at-first-survey by entry-year fixed; column (2) adds all the demographic, education, cognitive and non-cognitive skills[10] and ability controls to the regression, column (3) further adds all the

[10] The regression for the 1979 cohort separately includes measures of non-cognitive skills, in addition to cognitive skills

family background controls, and lastly column (4) includes all control variables[11] and initial region fixed effects. Standard errors are clustered by initial industry.

This positive association between industry employment growth and real earnings growth remains significant and robust, even after adding all individual characteristics and measures of education and ability to the baseline specification; and also, after adding the individual's family and parental background characteristics; and furthermore, after controlling for the initial region of residence. For example, the estimates for the 1979 cohort in column (4) suggest that a one standard deviation increase in the industry employment growth is associated with a 0.09 standard deviation increase in the individual's earnings growth over 10 years from the initial industry choice.

Thus, we find that the estimated effects of industry employment growth on earnings growth remains stable and significant even after controlling for individual ability, family background or initial region of residence. We find no effect on the estimates from family background or regional controls after controlling for individual attributes, except for the 1966 cohort to some extent[12]. We do not add other variables common in standard Mincerian-type wage functions, such as marital status, on-the-job experience, spousal variables, etc., because they are not relevant for answering our research question, which is essentially about the initial industry choice, and the pre-labor market endowment of an individual.

V.B: Robustness - Industry reported *immediately* after starting full-time employment

In the main results shown in section (VA), we consider within the sample of analysis all individuals who declare an initial industry within 7 years of starting full-time employment. We do this in order to gain the largest sample for our analysis.

However, one might argue that the initial industry should be considered as that which the individual reports being in, almost immediately after entering the labor market. In order to check the robustness of the previous results, we now restrict our sample to only those individuals who declared an industry within 2 years of starting full-time employment.

[11] Table A.2 in the appendix describes all of the control variables which were used for the three separate NLSY cohorts based on data availability

[12] This could be due to data constraints since for the 1966 cohort we have fewer detailed individual-level control variables.

Table 3: Growth in earnings 10 years after industry entry – industry reported within 2 yrs. of FT employment

Growth in (log) average real earnings	**(1)**	**(2)**	**(3)**	**(4)**
1966 cohort:				
Growth in industry employment	0.776***	0.361***	0.209*	0.188*
	(0.153)	(0.119)	(0.105)	(0.103)
Observations	2,888	2,888	2,888	2,888
R-squared	0.286	0.391	0.465	0.466
1979 cohort:				
Growth in industry employment	0.832***	0.522***	0.516***	0.511***
	(0.153)	(0.113)	(0.111)	(0.112)
Observations	4,071	4,071	4,071	4,071
R-squared	0.107	0.230	0.242	0.246
1997 cohort:				
Growth in industry employment	0.987***	0.609***	0.562***	0.579***
	(0.249)	(0.141)	(0.138)	(0.140)
Observations	1,581	1,581	1,581	1,581
R-squared	0.142	0.269	0.275	0.277
Age-at-first-survey x Entry-Year FE	YES	YES	YES	YES
All demographic + education variables		YES	YES	YES
All family background variables			YES	YES
Initial Region FE				YES

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

NOTE: Each panel represents separate OLS regressions (standard errors in parentheses clustered by initial industry) for the NLSY cohort. Data is aggregated up to time-consistent two-digit industries. Here, for each cohort, sample size is smaller than in Table 2 because we remove individuals who did not report industry within two years of full-time employment.

Once again, we find a positive and statistically significant association between the industry employment growth and the 10-year and 20-year growth in earnings of the individual, and the estimated coefficient on the industry growth variable remains significant and stable even after the addition of all control variables and fixed effects.

V.C: Robustness - Identifying initial industry after completion of schooling

In order to further assess the robustness of our findings, we now change the way we identify the individual's initial industry.

In the main analysis, we had identified initial industry as the one declared after the individual began full-time employment and we do not exclude anyone who is also undertaking schooling while being in full-time employment. Now, we change our identification of initial industry as that which the individual reports being employed in, after completing full-time schooling. This is a more restrictive way of identifying initial industry.

Table 4: Growth in earnings 10 years after industry entry – industry identified after completion of schooling

Growth in (log) average real earnings	**(1)**	**(2)**	**(3)**	**(4)**
1966 cohort:				
Growth in industry employment	0.512***	0.378***	0.214*	0.206*
	(0.118)	(0.121)	(0.115)	(0.110)
Observations	2,584	2,584	2,584	2,584
R-squared	0.177	0.200	0.262	0.263
1979 cohort:				
Growth in industry employment	0.508***	0.465***	0.463***	0.463***
	(0.145)	(0.148)	(0.140)	(0.139)
Observations	3,985	3,985	3,985	3,985
R-squared	0.063	0.085	0.094	0.102
1997 cohort:				
Growth in industry employment	0.785***	0.707***	0.719***	0.697***
	(0.239)	(0.213)	(0.235)	(0.229)
Observations	1,455	1,455	1,455	1,455
R-squared	0.095	0.138	0.151	0.154
Age-at-first-survey x Entry-Year FE	YES	YES	YES	YES
All demographic + education variables		YES	YES	YES
All family background variables			YES	YES
Initial Region FE				YES

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

NOTE: Each panel represents separate OLS regressions (standard errors in parentheses clustered by initial industry) for the NLSY cohort. Data is aggregated up to time-consistent two-digit industries. Here, for each cohort, sample size is different than in Table 2 because we identify the entry-year and the initial industry differently: we follow each individual till they complete full-time schooling and then consider the industry declared thereafter as the initial industry.

Our results in this robustness check are similar to our main results and also to the previous robustness check – thus providing validation to the results in our analysis. We find that the relationship between growth in earnings and the corresponding employment growth in the initial industry is positive and statistically significant, and the results remain stable even after the addition of all controls and fixed effects.

Thus, our main results and checks on robustness show a significant, positive and persistent relationship between long-term earnings growth and the long-term performance of the initial industry (in terms of employment growth). This relationship does not get eroded even after we control for the pre-market endowment that each individual brings into the labor market, and thus, it can be said that earnings growth is not just explained by the individual and the family background – a part of it also depends on the choice of industry made during entry into the labor market.

V.D: Main results - Explaining growth in earnings 20 years after entry year

Now we conduct the same analysis but for a longer period of time – 20 years from entry into the industry. Once again, we first provide results for the baseline specification, and then the results for the different specifications where we control for key individual characteristics, family or parental background characteristics, and then add fixed effects for initial region of residence. *However, due to data constraints from the 1966 and the 1997 cohorts, these results only take into account the 1979 cohort.*

Table 5: Growth in earnings 20 years after industry entry

Growth in (log) average real earnings	(1)	(2)	(3)	(4)
Growth in industry employment	0.404***	0.254***	0.244***	0.248***
	(0.0929)	(0.0702)	(0.0644)	(0.0649)
Constant	2.710***	1.893***	1.802***	1.972***
	(0.0476)	(0.287)	(0.401)	(0.455)
Observations	3,319	3,319	3,319	3,319
R-squared	0.156	0.328	0.347	0.349
Age-at-first-survey x Entry-Year FE	YES	YES	YES	YES
All demographic + education variables		YES	YES	YES
All family background variables			YES	YES
Initial Region FE				YES

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

NOTE: Table represents OLS regressions (standard errors in parentheses clustered by initial industry) for the NLSY 79 cohort. Data is aggregated up to time-consistent two-digit industries.

For 20 years after the individual's entry into the industry, the positive and statistically significant relationship between growth in earnings and growth in industry employment, once again, remains in all four specifications. Results from column (4) suggest that a one standard deviation increase in the industry employment growth rate over 20 years is associated with a 0.092 standard deviation increase in the individual's growth in earnings.

Thus, entering into a "favourable" industry, i.e., an industry in which the employment grew over time, has a significant positive influence on the earnings growth of an individual, and this influence remains even after controlling for initial region of residence, any perceived measures of ability and schooling and other important family background and demographic characteristics. This positive and significant influence is also persistent – it remains 20 years after the initial industry entry.

V.E: Heterogeneity Analysis[13]

(V.E.1): Heterogeneity Analysis – Worker persistence in initial industry

The relative importance of the initial industry would be diluted if the results differ according to whether a person stays in the industry for long or not. One might argue that the "*choice*" of the industry is less important compared to the "*length of stay*" within the industry. Thus, it would be interesting to see if the earnings growth is explained different by the growth in industry employment, based on worker persistence within the initial industry.

Table 6: Explaining growth in earnings by worker persistence in initial industry

Growth in (log) average real earnings (20 yrs.)	(1) Worker persisted for at least 5 years	(2) Worker did not persist for 5 years
Growth in industry employment (20 yrs.)	0.330***	0.211***
	(0.0890)	(0.0691)
Constant	0.0875	1.544***
	(0.919)	(0.520)
Observations	836	2,483
R-squared	0.420	0.354
All demographic + education variables	YES	YES
All family background variables	YES	YES
Age-at-first-survey x Entry-Year FE	YES	YES
Initial Region FE	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

NOTE: Table represents OLS regressions (standard errors in parentheses clustered by initial industry) for the NLSY 79 cohort. Data is aggregated up to time-consistent two-digit industries. In column (1) each individual reported being in the initial industry 5 years after entry-year.

We split our sample into those who left the initial industry within 4-5 years of entry, and those who persisted for at least 5 years. The results suggest that the significant and positive association between industry employment growth and earnings growth does not depend upon persistence within the initial industry or the amount of time spent in it. We find no significant difference in the estimates between those who persisted in their initial industries versus those who did not. The positive and significant association also holds for the sample that did not persist within the initial industry for even 4-5 years – suggesting that this early-career choice may have a significant association on long-term earnings trajectories, even if the worker decides to leave the industry after a short period of time.

(V.E.2): Heterogeneity Analysis – Task content at initial occupation

Another important factor explaining differential returns to an initial industry choice could be the skills needed or the tasks required to perform at the individual's initial occupation

[13] The heterogeneity analyses show results with earnings growth over 20 years as the outcome of interest, and only uses the 1979 NLSY cohort, due to data constraints for the other two cohorts.

within the industry. A considerable part of the literature on earnings inequality and skill-biased technological change has focused on the differential returns to occupational skills, and how that has driven the increasing inequality in earnings in the labor market (Autor et al., 2003; Acemoglu & Autor, 2011; Autor & Dorn, 2013). A large section of this literature has also documented the widespread decline of routine employment in the U.S.

Thus, in order to conduct this analysis, we follow the premise behind Autor & Dorn (2013) and split the sample into three groups based on individuals entering into occupations which have differential task content – above or below median manual, routine and abstract task content.

Table 7: Explaining growth in earnings by tasks in initial occupation

Growth in (log) average real earnings (20 yrs)	(1) Above median abstract	(2) Above median routine	(3) Above median manual
Growth in industry employment (20 yrs)	0.168	0.233**	0.320***
	(0.121)	(0.0847)	(0.0959)
Constant	-0.141	0.537	2.181***
	(0.331)	(0.579)	(0.492)
Observations	1,657	1,610	1,506
R-squared	0.406	0.361	0.415
Age-at-first-survey x Entry-Year FE	YES	YES	YES
All demographic + education variables	YES	YES	YES
All family background variables	YES	YES	YES
Initial Region FE	YES	YES	YES

Growth in (log) average real earnings (20 yrs)	(1) Below median abstract	(2) Below median routine	(3) Below median manual
Growth in industry employment (20 yrs)	0.328***	0.143	0.169*
	(0.0643)	(0.0925)	(0.0862)
Constant	1.913***	1.170***	0.823
	(0.616)	(0.413)	(0.563)
Observations	1,662	1,709	1,813
R-squared	0.356	0.391	0.345
Age-at-first-survey x Entry-Year FE	YES	YES	YES
All demographic + education variables	YES	YES	YES
All family background variables	YES	YES	YES
Initial Region FE	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

NOTE: Table represents OLS regressions (standard errors in parentheses clustered by initial industry) for the NLSY 79 cohort. Data is aggregated up to time-consistent two-digit industries. Occupational task content is measured on a scale of 0 to 10, with an increasing score representing an increase in the specific task content, and is based on the data from the Dictionary of Occupational Titles 1977 (Autor & Dorn, 2013). For this sample of analysis, the median score for abstract task content = 1.858; for routine task content = 3.715; for manual task content = 0.6574.

We find the effects to be concentrated among those workers who entered into occupations that required more manual or routine and fewer abstract tasks to be performed. Even though the differences in the estimates for the different groups are not statistically

significant, these results are very suggestive of certain patterns that have already been documented in the literature. These results are consistent with the hypothesis that individuals who performed relatively manual (low-skilled) or relatively routine tasks would be the ones more affected from any shocks to the industry they are employed in – be it positive or negative shocks - and as a result would be more affected by changes in industry employment. High-skilled individuals needing to perform more cognitive or more abstract tasks do seem to get positive returns, but they do not seem to be significantly responsive to their initial industry choices.

(V.E.3): Heterogeneity Analysis – Offshorability factor of occupation at initial industry

The issue of "off-shoring" of jobs and occupations within the United States to economies with lower wages has received considerable importance over the years by researchers and also policymakers. Outsourcing or offshoring of tasks has also been cited as one of the reasons for the increasing wage inequality in developed economies and has been linked to moderate wage reduction for low-skilled workers using panel data (Baumgarten et al., 2013). In this context, it would be interesting to investigate if the offshorability factor of the occupation entered into within the initial industry differentially explains the relationship between long-term earnings growth and industry employment growth.

Table 8: Explaining growth in earnings by offshorability factor of initial occupation

Growth in (log) average real earnings (20 yrs)	(1) Above-median offshorability	(2) Below-median offshorability
Growth in industry employment (20 yrs)	0.306**	0.192**
	(0.113)	(0.0781)
Constant	-0.954	2.020***
	(0.723)	(0.584)
Observations	1,689	1,627
R-squared	0.394	0.362
Age-at-first-survey x Entry-Year FE	YES	YES
All demographic + education variables	YES	YES
All family background variables	YES	YES
Initial Region FE	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

NOTE: Table represents OLS regressions (standard errors in parentheses clustered by initial industry) for the NLSY 79 cohort. Data is aggregated up to time-consistent two-digit industries. Offshorability is measured as a continuous variable, with an increasing score representing an increase in the specific task content, and is based on task values from O*NET (Autor & Dorn, 2013). For this sample of analysis, the median score for abstract task content = 1.858; for routine task content = 3.715; for manual task content = 0.6574.

Once again, we find that although the difference in the estimates between the two groups is not statistically significant, the effect is larger for the group of workers entering into

occupations which are easier to offshore. This is consistent with the growing literature on wage inequality that has found negative wage effects of offshoring. If the initial occupation has an above-median offshorability factor, making a "favorable" industry choice becomes all the more important for long-term growth in earnings.

Section VI: Conclusion

This paper is an empirical analysis of the relationship between a critical labor market choice made by individuals early in their careers and long-term growth in earnings. We find that making the choice of entering into a "good" or favorable industry, i.e., an industry which experienced high employment growth over time, leads to a significantly positive influence on a worker's long-term earnings growth. This association is also persistent over time – often lasting for 20 years from the individual's entry into the industry.

In order to control for the pre-labor market endowment that each individual possesses, we control for important individual and family background characteristics which may impact individual earnings, for any age effects, year of entry effects, any effects from the region of residence during entry into the industry. Including these controls does not erode or remove the significantly positive association that the industry choice has on long-term earnings growth, and the estimates remain stable even after the addition of several fixed effects.

Our findings are also robust to several checks – we change the way we identify an initial industry and we also apply the analysis to a smaller sample, and these robustness checks also find very results similar to the results in our main analysis. We also conduct a placebo test to assess the validity of the analysis[14], and this test does not find a relationship where one should not be.

From the heterogeneity analysis, we find that the exposure from the initial industry does not depend on how long a worker stays within it – the results remain positive and significant even for those who left the initial industry within 5 years of entry. Further analysis also suggests that the effects are concentrated around those who entered into occupations within the initial industry which involved above-median manual or routine tasks content – thus indicating that those who are employed in low-skill or manual occupations may be more sensitive to any "shocks" to the industry they are employed in – whether those shocks may be negative or positive. Looking into the offshorability-factor of the occupational task, we find that the effects are stronger for those individuals who entered into occupations which had an above-median offshorability factor, compared to those with below-median offshorability factor, even though we do not find statistically significant differences in the estimates for the two groups. However, these results overall suggest a pattern that the initial industry chosen by a young worker does matter considerably for longer-term outcomes by setting in place a trajectory which influences

[14] Results in the Appendix Table A.3

their earnings profiles and which their individual attributes, family background or regional characteristics cannot erode.

Appendix

A.1. Explaining initial (average) earnings by pre-market endowment:

In order to understand the sample of analysis even further, we attempt to show how the pre-market endowment of these individuals influence their average earnings at the time of entry into the initial industry.

Table A.1 below shows the relationship between the individual and family background characteristics (the pre-market endowment) and the average earnings of the first three years after initial entry into the industry. Since our dependent variables in the main analyses are the 10-year and 20-year *growth* in earnings from entry into the industry, Table A.1 shows how the individual control variables explain the level of earnings of a person accrued during entry into the industry.

Table A.1: Explaining initial earnings by pre-market endowment variables

Initial average real earnings	(1) NLSY 66	(2) NLSY 79	(3) NLSY 97
Black	-0.0979***	-0.139***	-0.199***
	(0.0276)	(0.0394)	(0.0553)
Hispanic	-0.117	0.0163	-0.0917
	(0.103)	(0.0360)	(0.0551)
Foreign born	-0.0816	-0.0793	-0.0190
	(0.0911)	(0.0584)	(0.0291)
Years of schooling	-0.00510	-0.0648***	-0.0271
	(0.00875)	(0.0141)	(0.0203)
High-school graduate	0.0465	0.145**	0.107**
	(0.0325)	(0.0526)	(0.0446)
Some college education	-0.0434	0.141**	0.0406
	(0.0434)	(0.0679)	(0.0625)
College graduate	-0.223**	0.0379	-0.165
	(0.0778)	(0.123)	(0.106)
Initial family income	0.409***	0.0603**	0.0715***
	(0.0400)	(0.0244)	(0.0246)
Father's years of schooling	0.00195	-0.00544	0.00243
	(0.00541)	(0.00427)	(0.0103)
Mother's years of schooling	0.0104	0.00551	-0.0259*
	(0.00634)	(0.00474)	(0.0150)
Father is an immigrant	0.0470	0.0945	0.160**
	(0.0797)	(0.0636)	(0.0648)
Mother is an immigrant	0.0708	0.0269	0.00521
	(0.0731)	(0.0637)	(0.0716)
Constant	4.209***	7.850***	6.460***
	(0.310)	(0.318)	(0.432)
Observations	3,002	4,308	1,638
R-squared	0.647	0.392	0.320
Age-at-first-survey x Entry-Year FE	YES	YES	YES
Initial Region FE	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

A.2: Variables used as controls in the 10-year analysis for the three NLSY cohorts

Variable name	Description	Used as control in NLSY 66	Used as control in NLSY 79	Used as control in NLSY 97
black	=1 if respondent is African-American	**Y**	**Y**	**Y**
Hispanic	=1 if the respondent is of Hispanic origin	N	**Y**	**Y**
other race	=1 if the respondent is of another race, not black or white	**Y**	N	N
born outside USA	=1 if respondent was not born within the United States	**Y**	**Y**	**Y**
yrs. of schooling	total years of schooling completed by respondent	**Y**	**Y**	**Y**
HS graduate	=1 if respondent only has a high school degree or a GED	**Y**	**Y**	**Y**
college graduate	=1 if respondent is a college graduate	**Y**	**Y**	**Y**
some college	=1 if respondent has completed some college, but has no degree	**Y**	**Y**	N
Composite IQ score	A composite score that combines the scores from various aptitude and intelligence tests which were collected from respondents; standardized scores are used	**Y**	N	N
all individual ASVAB components	10 different ASVAB components included; age-standardized scores used	N	**Y**	**Y**
Rotter locus of control (std.)	a measure of "non-cognitive" skills using the normalized average of the Rotter Locus of Control	N	**Y**	N
Self esteem score (std.)	a measure of "non-cognitive" skills using the normalized average of the Rosenberg Self-Esteem Scale	N	**Y**	N
sociability at age 6 (std.)	Self-reported sociability in 1981 at age 6 (retrospective) (extremely shy, somewhat shy, somewhat outgoing, extremely outgoing)	N	**Y**	N
sociability as adult (std.)	Self-reported sociability in 1981 (extremely shy, somewhat shy, somewhat outgoing, extremely outgoing)	N	**Y**	N
club participation	The number of clubs in which the respondent participated in high school	N	**Y**	N

initial family income	(log) of the family income during entry-year	**Y**	**Y**	**Y**
initial family below PL	=1 if family during entry-year was below the poverty line	N	**Y**	N
initial residence in SMSA	=1 if residence was within SMSA during entry-year	**Y**	**Y**	N
father's years of schooling	total years of schooling completed by respondent's father	**Y**	**Y**	**Y**
mother's years of schooling	total years of schooling completed by respondent's mother	**Y**	**Y**	**Y**
male in HH	=1 if the respondent grew up with a working male in the household	**Y**	**Y**	N
female in HH	=1 if respondent grew up with a working female in the household	**Y**	**Y**	N
father's occupation dummies	=1 if different occupation dummies included, based on father's main occupations reported	N	**Y**	N
same occ as father's	=1 if respondent entered into the same occupation as father's main occupation at entry-year	N	**Y**	N
foreign language spoken at home	=1 if respondent grew up speaking a non-English (foreign) language at home	**Y**	**Y**	N
father foreign-born	=1 if father was an immigrant	**Y**	**Y**	N
mother foreign-born	=1 if mother was an immigrant	**Y**	**Y**	N

A.3: Placebo Test

In order to test the validity of our results, and to check that our analysis does not show any association where there should not be any, we conduct a placebo test using the data from the 1979 cohort where we find the relationship between an individual's earnings acquired **in the first 10 years** from entry into the industry, and the growth in industry employment **after the first 10 years**, i.e., growth between 10 and 20 years from entry. Our hypothesis is that an individual's earnings in the first 10 years from entry into the industry should have no significant association with any industry growth taking place after those 10 years, since those years have not taken place yet.

Table A.3: Explaining real earnings acquired *in the first 10 years after entry*

Avg. real earnings: 10 years after entry	1	2
Industry employment growth in the next 10 years	-0.0463	-0.146
	(0.126)	(0.108)
Constant	7.018***	5.351***
	(0.0276)	(0.333)

Observations	4,261	4,261
R-squared	0.086	0.333
Age-at-first-survey x Entry-Year FE	YES	YES
All demographic + education variables		YES
All family background variables		YES
Initial Region FE		YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

From the placebo test, we find that there no significant association between earnings 10 years after entering the industry, and any employment growth that took place in the industry in the subsequent 10 years, and this non-association remains even when we add all control variables and all fixed effects.

References

Abowd, J. M., McKinney, K. L., & Zhao, N. L. (2018). Earnings inequality and mobility trends in the United States: Nationally representative estimates from longitudinally linked employer-employee data. *Journal of Labor Economics*, *36*(S1), S183-S300.

Acemoglu, D., & Autor, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. *Handbook of labor economics*, *4*, 1043-1171.

Altonji, J. G., Bharadwaj, P., & Lange, F. (2012). Changes in the characteristics of American youth: Implications for adult outcomes. *Journal of Labor Economics*, *30*(4), 783-828.

Altonji, J. G., Kahn, L. B., & Speer, J. D. (2016). Cashier or consultant? Entry labor market conditions, field of study, and career success. *Journal of Labor Economics*, *34*(S1), S361-S401.

Altonji, J. G., & Pierret, C. R. (2001). Employer learning and statistical discrimination. *The Quarterly Journal of Economics*, *116*(1), 313-350.

Arellano-Bover, J. (2019). Career Consequences of Firm Heterogeneity for Young Workers: First Job and Firm Size. *Job Market Paper*, Stanford University.

Åslund, O., & Rooth, D. O. (2007). Do when and where matter? Initial labour market conditions and immigrant earnings. *The Economic Journal*, *117*(518), 422-448.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., & Van Reenen, J. (2017). *The fall of the labor share and the rise of superstar firms*. National Bureau of Economic Research.

Autor, D. H. (2014). Skills, education, and the rise of earnings inequality among the "other 99 percent". *Science*, *344*, 843-851.

Autor, D. H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *The American Economic Review*, *103*(5), 1553-1597.

Autor, D. H., Dorn, D., & Hanson, G. H. (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review*, *103*(6), 2121-68.

Autor, D. H., Katz, L. F., & Kearney, M. S. (2008). Trends in US wage inequality: Revising the revisionists. *The Review of economics and statistics*, *90*(2), 300-323.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, *118*(4), 1279-1333.

Barth, E., Bryson, A., Davis, J. C., & Freeman, R. (2016). It's where you work: Increases in the dispersion of earnings across establishments and individuals in the United States. *Journal of Labor Economics*, *34*(S2), S67-S97.

Baumgarten, D., Geishecker, I., & Görg, H. (2013). Offshoring, tasks, and the skill-wage pattern. *European Economic Review*, *61*, 132-152.

Burgess, S., Propper, C., Rees, H., & Shearer, A. (2003). The class of 1981: the effects of early career unemployment on subsequent unemployment experiences. *Labour Economics*, *10*(3), 291-309.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. Journal of Labor Economics 36(S1), S13–S70.

Card, D., Heining, J., & Kline, P. (2013). Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly journal of economics*, *128*(3), 967-1015.

Couch, K. A., & Placzek, D. W. (2010). Earnings losses of displaced workers revisited. *The American Economic Review*, *100*(1), 572-589.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The Quarterly Journal of Economics*, *132*(4), 1593-1640.

Dickens, W., & Katz, L. F. (1986). Interindustry wage differences and industry characteristics.

Dickens, W., & Katz, L. F. (1987). Inter-industry wage differences and theories of wage determination.

Gaini, M., Leduc, A., & Vicard, A. (2012). *A scarred generation? French evidence on young people entering into a tough labour market* (No. g2012-05). Institut National de la Statistique et des Etudes Economiques, DESE.

Gardecki, R., & Neumark, D. (1998). Order from chaos? The effects of early labor market experiences on adult labor market outcomes. *ILR Review*, *51*(2), 299-322.

Hanushek, E. A., Schwerdt, G., Wiederhold, S., & Woessmann, L. (2015). Returns to skills around the world: Evidence from PIAAC. *European Economic Review*, *73*, 103-130.

Kahn, L. B. (2010). The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics*, *17*(2), 303-316.

Katz, L. F., & Autor, D. H. (1999). Changes in the wage structure and earnings inequality. *Handbook of Labor Economics*, *3*, 1463-1555.

Kopczuk, W., Saez, E., & Song, J. (2010). Earnings inequality and mobility in the United States: evidence from social security data since 1937. *The Quarterly Journal of Economics*, *125*(1), 91-128.

Krueger, A. B., & Summers, L. H. (1988). Efficiency wages and the inter-industry wage structure. *Econometrica: Journal of the Econometric Society*, 259-293.

Kyui, N., & Simonnet, V. (2013). Education, Professional Choice and Labour Market Outcomes: Influence of Preferences, Parental Background and Labour Market Tightness.

Lemieux, T. (2006). Increasing residual wage inequality: Composition effects, noisy data, or rising demand for skill?. *The American Economic Review*, *96*(3), 461-498.

Liu, K., Salvanes, K. G., & Sørensen, E. Ø. (2016). Good skills in bad times: Cyclical skill mismatch and the long-term effects of graduating in a recession. *European Economic Review*, *84*, 3-17.

Muller, S. and R. Neubaeumer (2018). Size of training firms: The role of firms, luck, and ability in young workers' careers. International Journal of Manpower 39(5), 658–673.

Mueller, H. M., Ouimet, P. P., & Simintzi, E. (2017). Wage inequality and firm growth. *American Economic Review*, *107*(5), 379-83.

Oreopoulos, P., Von Wachter, T., & Heisz, A. (2012). The short-and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics*, *4*(1), 1-29.

Oyer, P. (2006). Initial labor market conditions and long-term outcomes for economists. *The Journal of Economic Perspectives*, *20*(3), 143-160.

Oyer, P. (2008). The making of an investment banker: Stock market shocks, career choice, and lifetime income. *The Journal of Finance*, *63*(6), 2601-2628.

Piketty, T., & Saez, E. (2003). Income inequality in the United States, 1913–1998. *The Quarterly journal of economics*, *118*(1), 1-41.

Raaum, O., & Røed, K. (2006). Do business cycle conditions at the time of labor market entry affect future employment prospects?. *The review of economics and statistics*, *88*(2), 193-210.

Schwandt, H. and T. von Wachter (2017). Unlucky Cohorts: Earnings, Income, and Mortality Effects from Entering the Labor Market in a Recession. Mimeo (April 2017).

Spletzer, J. R. (2014, June). Inequality Statistics from the LEHD. In *June 2014 FESAC meetings, http://www. census. gov/fesac/pdf/2014-06-13/Spletzer_Background. pdf*.

von Wachter, T., & Bender, S. (2006). In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers' Careers. *American Economic Review*, *96*(5), 1679-1705.