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 long term earnings growth, while conditioning on a large number of controls including 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 (1966, 1979 and 1998) and historical industry data from the Bureau of Labor Statistics. The results suggest that there is a significant, positive and persistent relationship between earnings growth and making a "favorable" industry choice in terms of future industry employment growth. The largest effects are observed for 1979 cohort entering the labor market in the 1980's consistent with a period of growing concerns about job loss due to automation and international trade. Surprisingly, however, we only find evidence that these effects are concentrated among occupation types which involve more routine tasks and easier off-shorability for the 1966 cohort.

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

Industries have long been characterized by differences in wages that cannot be explained by observed worker attributes (Dickens and Katz 1986, 1987; Krueger and Summers 1988). Industries may provide different opportunities to obtain industry specific skills (Neal 1995; Parent 2000; Kwon and Milgram 2014). Industries vary dramatically in their size distribution, and large firms appear to pay a wage premium over smaller firms (Lochner et al. 2020; Bloom et al. 2018). Mueller, Ouimet and Simintzi (2017) also document that larger firms have more pay inequality. Over time, industries where production involves more routine tasks have also seen significant changes in labor demand that have likely influenced wages (Autor, Levy and Murnane 2003). These and other differences across firms and industries may have substantial impacts on the life-time earnings of workers in these industries.

Research has also documented that the long-run earnings of workers depend upon the circumstances when they first enter the labor market and on the attributes of their first job. A wide array of studies (Åslund and Rooth 2007; Kahn 2010; Liu, Salvanes and Sørensen 2016; Oreopoulos, Von Wachter and Heisz 2012; Oyer 2008, 2009; Raaum and Røed 2006; Schwandt and von Wachter 2017, 2019) find that entering into the labor market during a recession has long lasting, negative impact on worker earnings.[1] Altonji, Kahn and Speer (2016), Oreopoulos, Von Wachter and Heisz (2012), and Schwandt and von Wachter (2017) find that effects are larger and/or more long-lasting for less skilled or more disadvantaged workers. Arellano-Bover (2019) and Muller and Neubaeumer (2018) find positive and persistent effects of early employment at a large firm or establishment on income and employment, respectively. Also, Abowd, McKinney, and Zhao (2018) document large and persistent wage gains from working at the highest paying firms at all skill levels.

[1] Exceptions are Altonji, Kahn and Speer (2016) and Gaini, Leduc and Vicard (2012) who find that these effects fade relatively quickly.

In this study, we examine whether an individual's initial industry has a long-run impact on earnings through one particular lens. Specifically, we test whether future employment growth of the industry in which an individual first works influences their future earnings. Several studies document that large job displacements can have substantial and long lasting impacts on worker wages and earnings (Burgess, Propper, Rees, and Shearer 2003; Couch and Placzek 2010; Jacobson, LaLonde, and Sullivan 1993), and Burgess, Propper, Rees, and Shearer (2003) document more persistent effects for unskilled workers. On the other hand, Gardecki and Neumark (1998) and von Wachter and Bender (2006) find that job losses early in ones work career have at most modest and short-lived effects on earnings. Further, the importance of initial industry choice may have changed over time given the influence of technology on the prospects of low skill workers employed in routine tasks (Acemoglu and Autor 2011; Autor and Dorn 2013).

In order to address these questions, we examine the earnings growth of male workers using data from the National Longitudinal Survey's of Young Men/Youth initially conducted in 1966, 1979 and 1997. A natural concern about such an analysis is that workers may sort into industries based on their unobserved ability. For example, Abowd, Kramarz and Margolis (1999) find that 75% of the wage premium associated with large firms and 90% of industry wage differentials can be explained by those firms hiring higher quality workers. A unique advantage of these three National Longitudinal Surveys is that they conduct a battery of aptitude tests (Armed Services Vocational Aptitude Battery), for the last two surveys collect data on non-cognitive skills, and also provide detailed information on the family background of the youth, all of which may capture the unobserved productivity of these young workers.

In all three surveys, we find that respondents selecting an initial industry that has negative or lower employment growth over the next 10 years experience lower earnings growth. The estimated effects of initial industry are largest for the 1979 cohort with a one standard deviation decline in employment growth implying 9 percent of a standard deviation lower wage growth. The standardized

effect for the 1966 cohort is only 66% of the effect for 1979. The estimated effect in 1997 is the larger than the 1979 estimate, but the standardized effect is 15 percent smaller because the volatility in employment growth across industries fell by 25 percent between the 1979 and 1997 cohorts. After controlling for individual demographics, the results of the aptitude battery tests, and non-cognitive skills for the 1979 and 1997 cohorts, the inclusion of additional controls for family background or for region of the country have minimal impact of our estimates. This result is consistent with the individual controls capturing a substantial share of unobserved earnings capacity. For the 1979 cohort, we document a similar effect over twenty years of earnings growth, and also show in a placebo test that employment growth between 10 and 20 years after initial industry choice (future growth) cannot explain wage growth during the first 10 years after initial industry choice.

Next, we examine whether these effects are heterogeneous across workers. We first demonstrate that the effects on wage growth are similar when comparing subsamples of workers who left the industry during their first four years after entering the labor market and those who remained in the industry for five or more years. For the 1966 cohort, we also demonstrate that the effects on wage growth are larger for workers whose initial occupations involve above median levels of routine tasks, and score more highly on a measure of off-shorability. However, we do not observe any evidence of this heterogeneity for the 1979 and 1997 cohorts. If anything, the estimates while noisy suggest that by the 1997 cohort the effects were larger for workers in occupations with below median share of routine tasks. In summary, the economic success of the industry initially chosen by a worker when they enter the labor market appears to impact earnings 10 and even 20 years after this choice, and these effects have been present over the last forty years.

Section II: 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 choice 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 a favourable choice. 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 initial full-time entry in the labor market as captured by working hours. We identify the first NLSY survey year when an individual is in full time employment (30 hours per week or more) as initial full-time entry, In order to maximize the sample, we consider all individuals who have declared an industry within 7 years of starting full-time employment. Earnings growth is then calculated from the first year that an individual reports being employed in a specific industry.

Our estimating equation is cross-sectional at the individual level and takes the following form

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

where Y_i is the outcome of interest (growth in earnings) pertaining to individual i; G_i refers to the employment growth in individual i's initial industry and β_1 is our coefficient of interest; 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.

Section III: Data & Summary Statistics

III.A: Data

The data for this paper comes from three different survey cohorts of the National Longitudinal Survey of Youth – the 1966 Young Men cohort (part of the NLSY Original Cohort

[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.

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 initially employed in the agricultural sector or the armed forces.[3] To obtain our sample of analysis: we first follow individuals over time to observe the year they declare full-time employment, which we define as their entry year.[4] However, some employed individuals do not report an industry especially in their early years in the labor market. In order to maximize sample, we include all individuals in our sample who declare an industry of employment within 7 years of first reporting full-time employment. Accordingly, we base a workers initial industry entry year as the first year in which they report an industry of employment following their initial report of full time employment in the labor market.

For the purposes of calculating earnings growth, we calculate earnings growth from the year that they first declare an industry following their labor market entry year. We base earnings growth on three-survey moving average of earnings around their initial industry choice year and they 10 or 20 years later in order to have the potential of at least three years of earnings data for

[3] 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.
[4] Full-time employment is defined as working 30 or more hours per week

each average, and we remove those survey year in which earnings are either reported as zero or earnings are not reported for any of the years included in the moving average.[5] 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.

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. Parental education and family income are included to capture information on the youth's family background. Therefore, if the youth is living with his/her parents or guardian at the time of the wave 1 survey, family income is set to zero and dummy variable for family income not available is included in the model and interacted with the parental education variables in order to allow those variables to be more important when family income is not observed. Further, family income excludes the earnings of the youth themselves in wave 1 because our dependent variable is based on youth's current or future earnings. 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] Specifically, for the 1979 cohort we include the Rosenberg self-esteem score and the Rotter Locus

[5] When surveys are conducted every two years, our moving average potentially covers a five year period.
[6] We have detailed 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.

of Control score (both standardized), self-reported sociability and participation in clubs. For the 1997 cohort, we include (standardized) measures of the respondent's self-reported scores on three questions in the survey that relate to the Big Five Personality Trait – Conscientiousness, Openness and Extraversion.[7]

III.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. Men who entered into industries which expanded over 10 years have higher perceived levels of education and ability and also parental education. In the 1966 cohort, differences in the sample across industries are relatively modest. However, in the 1979 and 1997 cohorts, we observed substantially higher own education and parental education levels in expanding industries, although differences in family income between industries are modest except in 1966. In all three cohorts, the surveys indicate substantially higher earnings growth in expanding industries in the 10 years following initial industry choice.

Section V: Results

IV.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. 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 effects; column (2) adds all the demographic, education,

[7] Description of these variables are provided in the Appendix Table A.2.

cognitive and non-cognitive skills[8] and ability (individual) controls to the regression, column (3) further adds all the family background controls, and lastly column (4) includes all control variables[9] and initial region fixed effects. Standard errors are clustered by initial industry.

While the inclusion of individual controls substantially reduces the estimates consistent with positive selection into industries that will experience future employment growth, additional controls for family background or region do not erode the estimates. In terms of magnitudes, the estimates for the 1979 cohort in column (4) suggest that a one standard deviation increase in the initial industry employment growth is associated with a 0.087 standard deviation increase in the individual's earnings growth over 10 years. The estimated effect for the 1966 cohorts is only 66% of the estimate for the 1979 cohort. The estimated effect for the 1997 cohort is larger than the estimate for 1979, but the standardized effect is 15 percent smaller because the volatility in employment growth across industries fell by 25 percent.

We next conduct two robustness tests related to our definition of initial industry. Our initial sample was based on all individuals who declare an initial industry within 7 years of starting full-time employment. 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 this alternative model, we now restrict our sample to only those individuals who declared an industry within 2 years of starting full-time employment. These results are show in Table 3. Once again, we find a positive and statistically significant association between the industry employment growth and the magnitude of the estimates are similar to those in Table 2 with a modest increase in the effect for the 1997 cohort. We also identify an individual's initial industry based on the first

[8] The regression for the 1979 cohort separately includes measures of non-cognitive skills, in addition to cognitive skills
[9] 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

industry reported after the individual has completed/left school. These results are shown in Table 3. Again, all results are robust and changes in magnitude are modest, less than 10 percent. Similarly, Table 4 presents results where initial industry is the first industry reported after has completed their initial spell of continuous education. Results are very similar except for the 1997 where the estimated effects are between 15 and 20 percent larger when initial industry measured after completing education.

IV.B: Main results - Explaining growth in earnings 20 years after entry year

Now, we exploit the longer time frame available only for the 1979 NLSY in order to conduct the same analysis but over 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. These results are shown in Table 5. Again, estimated effects are eroded substantially by the inclusion of individual controls, but are relatively stable as we include controls for family background and region. The effects are approximately one-half the magnitude of the effects over 10 years. However, the standard deviation of employment growth has also increased over the longer time period, and as a result the standardized effect of initial industry choice is virtually unchanged from the 10 year time horizon at 0.092.

The longer time horizon available for the NLSY 1979 allows us to conduct a simple falsification or placebo test where we examine whether future employment growth in the initial industry can predict past earnings growth. Specifically, we test for a relationship between an individual's earnings growth over the first 10 years after entry into the industry and the future employment growth of this industry employment, i.e., growth between 10 and 20 years after

industry 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 6 presents these results and the estimates are statistically insignificant and modest in magnitude.

V.C: Heterogeneity Analysis

In this subsection, we split our sample along several key dimensions to examine heterogeneity in the effects of initial industry choice. For example, in a world where industry specific human capital is important, the "*choice*" of the industry may be less important compared to the "*length of stay*" within the industry. Thus, we split our sample into those who left the initial industry within 4 years of entry, and those who persisted for at least 5 years. Table 7 presents these results for the 1997 cohort. Based on results in Table 7, the estimates are quite similar between the two subsamples, and we cannot reject the null hypothesis that the two estimates are the same.

Next, in Table 8, we present results dividing occupations based on a measure of off-shorability (Baumgarten et al., 2013). In this case, we find large differences with the estimates for above median off-shorability for the 1966 cohort being 36% larger than the comparable estimates in Table 2, and the estimates for below median off-shorability are near zero. The difference between the estimates has a T-statistic of 2.07. We find substantially smaller and statitistically insignificant differences between the subsamples for the 1979 and 1997 cohorts. Similarly, in Table 9, we split the samples based on the task content of the workers occupation within their initial industry dividing the occupations into above and below median on share of tasks involving abstract content, routine activities, and manual labor, e.g. Autor et al. (2003), Acemoglu and Autor (2011) and Autor & Dorn (2013). Again for the 1966 cohort, we find effects only for the subsample that has an above median share of routine tasks, and the t-statistics associated with this difference is 2.80. As with off-shorability, we do not find any statistically significant differences between

subsamples for the 1979 and 1997 cohort. However, focusing again on share of routine tasks, the subsample experiencing the largest effects of initial industry choice based on the point estimates has reversed with the effect for below median share on routine tasks more than double the effect for the above share subsample, but the associated t-statistic is only 1.09.

Section V: Conclusion

This paper presents an empirical analysis of the relationship between a critical labor market choice or outcome of individuals early in their career, initial industry, and their long-term growth in earnings. We find that making the choice of entering into a "good" or favorable industry, i.e., an industry which after the fact 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 including a variety of measures of cognitive and non-cognitive ability, 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 between success of the industry selected and long-term earnings growth. Our findings are also robust to several checks changing the way we identify an initial industry and finding very similar results. We also conduct a placebo test to assess the validity of the analysis showing that employment growth between years 10 and 20 has no impact on earnings growth during the first 10 years after initial industry choice.

We also investigate whether these effects are heterogeneous. We only find strong evidence of heterogeneity for the 1966 cohort. Specifically, we find that individuals working in occupations

with higher levels of off-shorability or higher shares of routine tasks experience substantially larger effects of initial industry choice consistent with workers engaged in routine activities facing threats from automation and from international trade. None of the differences across subsamples are significant for the other two cohorts. However, by the 1997 cohort, the relative magnitudes of the estimated effects between above and below median share of routine tasks appears to have reversed consistent with the effects of initial industry choice having greater impacts over time on workers doing non-routine tasks.

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.

Bloom, N., Guvenen, F., Smith, B.S., Song, J., & von Wachter, T. 2018. The disappearing large-firm wage premium. *AEA Papers and Proceedings*, 108, 317-322.

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.

Colonnelli, E., Tåg, J., Webb, M., & Wolter, S. 2018. A cross-country comparison of dynamics in the large firm wage premium. *AEA Papers and Proceedings*, 108, 323-327.

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.

Kwon, I., & Milgrom, E.M.M. 2014. The significance of firm and occupation specific human capital for hiring and promotions. *Labour Economics*, 31, 162-173.

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.

Lochner, B., Seth, S., & Wolter, S. 2020. Decomposing the large firm wage premium in Germany. *Economics Letters*, 194, 109368.

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.

Neal, D. 1995. Industry-Specific Human Capital: Evidence from Displaced Workers. *Journal of Labor Economics*, 13(4), 653–677

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.

Parent, D. 2000. Industry-Specific Capital and the Wage Profile: Evidence from the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics. *Journal of Labor Economics*, 18(2), 306–323.

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.

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
	(0.008)	(0.039)	(0.040)
Other race[10]	0.007	0.00	0.0065
	(0. 0015)	(0.00)	(0.001)
Age at initial survey	18.47	17.05	1.413***
	(0.059)	(0.247)	(0.296)
Born outside the U.S.	0.023	0.00	0.023**
	(0.003)	(0.00)	(0.014)
Total years of schooling	13.25	13.07	0.184
	(0.051)	(0.264)	(0.258)
College graduate	0.26	0.23	0.0295
	(0.008)	(0.039)	(0.0416)
High school graduate	0.34	0.34	-0.0077
	(0.008)	(0.044)	(0.0448)
Some college education	0.23	0.28	-0.0423
	(0.008)	(0.042)	(0.0402)
Father's years of education	1.78	0.551	1.231***
	(0.074)	(0.226)	(0.371)
Mother's years of education	2.158	0.526	1.632***
	(0.082)	(0.214)	(1.311)
Initial family income ($)	27.05	62.54	-35.494**
	(3.904)	(25.809)	(20.149)
Initial region: Non-south	0.60	0.59	0.0153
	(0.009)	(0.045)	(0.046)
Growth in (log) average real earnings (over 10 years)	1.028	0.881	0.147
	(0.018)	(0.094)	(0.091)
Growth in industry employment (over 10 years)	0.222	-0.059	0.281***
	(0.003)	(0.008)	(0.014)
1979 cohort			
Black	0.28	0.22	0.060***
	(0.007)	(0.013)	(0.016)
Hispanic	0.18	0.19	-0.008
	(0.006)	(0.013)	(0.014)
Age at initial survey	17.51	18.17	-0.657***
	(0.038)	(0.076)	(0.084)
Born outside the U.S.	0.0685	0.088	-0.0196***
	(0.044)	(0.083)	(0.009)
Total years of schooling	13.29	12.55	0.738***
	(0.044)	(0.044)	(0.096)

[10] The 1966 survey does not classify "Hispanic" as a separate race, while the 1979 and the 1997 surveys do.

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	9.671 (0.092)	8.967 (0.170)	0.703*** (0.161)
Mother's years of education	10.347 (0.0713)	9.582 (0. 1385)	0.764*** (0.126)
Initial family income ($)	10094.82 (218.692)	9772.008 (416.188)	416.188 (477.0868)
Initial region: North-East	0.19 (0.006)	0.16 (0.012)	0.030*** (0.014)
Growth in (log) average real earnings (over 10 years)	1.715 (0.0175)	1.440 (0.0294)	0.2745*** (0.0373)
Growth in industry employment (over 10 years)	0.229 (0.002)	-0.0859 (0.003)	0.3159*** (0.004)
1997 cohort			
Black	0.20 (0.012)	0.15 (0.014)	0.0439*** (0.019)
Hispanic	0.19 (0.012)	0.19 (0.015)	-0.0003 (0.019)
Age at initial survey	15.03 (0.044)	15.39 (0.050)	-0.367*** (0.068)
Born outside the U.S.	0.185 (0.012)	0.198 (0.015)	-0.0137 (0.0197)
Total years of schooling	14.02 (0.091)	13.57 (0.107)	0.458*** (0.142)
College graduate	0.35 (0.0152)	0.30 (0.0176)	0.0485*** (0.023)
High school graduate	0.45 (0.0159)	0.47 (0.019)	-0.028 (0.025)
Some college education	0.28 (0.014)	0.24 (0.016)	0.041** (0.022)
Father's years of education	11.58 (0.1634)	10.925 (0.197)	0.655*** (0.256)
Mother's years of education	12.55 (0.1186)	11.988 (0.144)	0.562*** (0.186)
Initial family income ($)	23756.2 (1412.366)	24269.8 (1521.594)	-513.601 (2119.007)
Initial region: North-East	0.17 (0. 012)	0.17 (0. 014)	0.002 (0.018)
Growth in (log) average real earnings (over 10 years)	1.563 (0.0389)	1.385 (0.0425)	0.177*** (0.0587)
Growth in industry employment (over 10 years)	0.0725 (0.002)	-0.1180 (0.004)	0.191*** (0.004)

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

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***	0.396***	0.375***
	(0.166)	(0.123)	(0.114)	(0.109)
Observations	3,002	3,002	3,002	3,002
R-squared	0.285	0.389	0.400	0.403
1979 cohort:				
Growth in industry employment	0.815***	0.513***	0.517***	0.515***
	(0.144)	(0.108)	(0.104)	(0.105)
Observations	4,305	4,305	4,305	4,305
R-squared	0.105	0.223	0.236	0.240
1997 cohort:				
Growth in industry employment	0.969***	0.576***	0.562***	0.584***
	(0.241)	(0.129)	(0.137)	(0.142)
Observations	1,638	1,638	1,638	1,638
R-squared	0.147	0.276	0.284	0.286
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.

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***	0.373***	0.348***
	(0.153)	(0.119)	(0.114)	(0.107)
Observations	2,888	2,888	2,888	2,888
R-squared	0.286	0.391	0.403	0.406
1979 cohort:				
Growth in industry employment	0.832***	0.522***	0.524***	0.521***
	(0.153)	(0.113)	(0.113)	(0.115)
Observations	4,071	4,071	4,071	4,071
R-squared	0.107	0.230	0.241	0.245
1997 cohort:				
Growth in industry employment	0.987***	0.620***	0.601***	0.620***
	(0.249)	(0.147)	(0.160)	(0.164)
Observations	1,581	1,581	1,581	1,581
R-squared	0.142	0.275	0.281	0.283
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.

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***	0.403***	0.390***
	(0.118)	(0.121)	(0.113)	(0.108)
Observations	2,584	2,584	2,584	2,584
R-squared	0.177	0.200	0.223	0.224
1979 cohort:				
Growth in industry employment	0.508***	0.465***	0.461***	0.460***
	(0.145)	(0.148)	(0.139)	(0.140)
Observations	3,985	3,985	3,985	3,985
R-squared	0.063	0.085	0.097	0.104
1997 cohort:				
Growth in industry employment	0.785***	0.712***	0.735***	0.714***
	(0.239)	(0.208)	(0.231)	(0.227)
Observations	1,464	1,455	1,455	1,455
R-squared	0.095	0.148	0.159	0.162
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.

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***	0.246***	0.248***
	(0.0929)	(0.0702)	(0.0692)	(0.0708)
Constant	2.710***	1.893***	1.641***	1.567***
	(0.0476)	(0.287)	(0.331)	(0.342)
Observations	3,319	3,319	3,319	3,319
R-squared	0.156	0.328	0.342	0.342
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.

Table 6: Placebo test for 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	0.00130	-0.111
	(0.127)	(0.124)
Constant	7.007***	6.734***
	(0.0279)	(0.256)
Observations	4,308	4,308
R-squared	0.085	0.341
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.

Table 7: Explaining growth in earnings by worker persistence in initial industry

	(1)	(2)
Growth in (log) average real earnings (10 yrs.)	Worker persisted for at least 5 years	Worker did not persist for 5 years
1966 cohort:		
Growth in industry employment (10 yrs.)	0.491***	0.336**
	(0.158)	(0.145)
Observations	979	2,023
R-squared	0.414	0.398
1979 cohort:		
Growth in industry employment (10 yrs.)	0.496***	0.510***
	(0.139)	(0.118)
Observations	1,071	3,234
R-squared	0.395	0.237
1997 cohort:		
Growth in industry employment (10 yrs.)	0.447	0.690***
	(0.403)	(0.212)
Observations	413	1,225
R-squared	0.329	0.306
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

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

	(1)	(2)
Growth in (log) average real earnings (10 yrs.)	Above median offshorability	Below median offshorability
1966 cohort:		
Growth in industry employment (10 yrs.)	0.509***	0.0680
	(0.117)	(0.178)
Observations	1,682	1,290
R-squared	0.421	0.428
1979 cohort:		
Growth in industry employment (10 yrs.)	0.479***	0.567***
	(0.110)	(0.162)
Observations	2,172	2,128
R-squared	0.274	0.258
1997 cohort:		
Growth in industry employment (10 yrs.)	0.614**	0.425*
	(0.263)	(0.214)
Observations	847	783
R-squared	0.315	0.340
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. 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.

Table 9: Explaining growth in earnings by tasks in initial occupation

1966 Cohort	Share Abstract	Share Routine	Share Manual
Growth in (log) average real earnings (10 yrs)	Above median	Above median	Above median
Growth in industry employment (10 yrs)	0.436***	0.633***	0.474***
	(0.144)	(0.123)	(0.121)
Observations	1,458	1,466	1,297
	Below median	Below median	Below median
Growth in industry employment (10 yrs)	0.457***	0.0485	0.320*
	(0.102)	(0.169)	(0.169)
Observations	1,544	1,536	1,705
1979 Cohort	Share Abstract	Share Routine	Share Manual
Growth in (log) average real earnings (10 yrs)	Above median	Above median	Above median
Growth in industry employment (10 yrs)	0.460**	0.379***	0.568***
	(0.180)	(0.131)	(0.172)
Observations	2,148	2,105	1,995
	Below median	Below median	Below median
Growth in industry employment (10 yrs)	0.580***	0.536***	0.454***
	(0.114)	(0.164)	(0.0952)
Observations	2,157	2,200	2,310
1997 Cohort	Share Abstract	Share Routine	Share Manual
Growth in (log) average real earnings (10 yrs)	Above median	Above median	Above median
Growth in industry employment (10 yrs)	0.731**	0.419	0.483**
	(0.297)	(0.280)	(0.230)
Observations	882	775	794
	Below median	Below median	Below median
Growth in industry employment (10 yrs)	0.565**	0.883**	0.520*
	(0.255)	(0.319)	(0.287)
Observations	756	863	844

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.

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

	(1)	(2)	(3)
Initial average real earnings	NLSY 66	NLSY 79	NLSY 97
Black	-0.242***	-0.300***	-0.284***
	(0.0260)	(0.0360)	(0.0429)
Hispanic /Other race	-0.130	-0.00214	-0.131**
	(0.105)	(0.0389)	(0.0520)
Foreign born	-0.00153	-0.0336	-0.0131
	(0.0614)	(0.0590)	(0.0301)
Years of schooling	-0.00481	-0.0467***	-0.0299
	(0.0118)	(0.0116)	(0.0183)
High-school graduate	0.0941**	0.273***	0.101**
	(0.0385)	(0.0495)	(0.0423)
Some college education	-0.0148	0.313***	0.0545
	(0.0581)	(0.0566)	(0.0594)
College graduate	-0.303**	0.198*	-0.129
	(0.106)	(0.114)	(0.112)
Initial family income	-0.00395	0.0183**	0.0448
	(0.0112)	(0.00812)	(0.0394)
Father's years of schooling	0.00731	0.000363	0.00326
	(0.00582)	(0.00523)	(0.0115)
Mother's years of schooling	0.0242**	0.0146***	-0.0203
	(0.00846)	(0.00472)	(0.0157)
Constant	6.623***	7.490***	7.060***
	(0.376)	(0.173)	(0.498)
Observations	3,002	4,308	1,638
R-squared	0.509	0.314	0.297
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