I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11114

Occupational Match Quality and Gender over Two Cohorts

John T. Addison
Liwen Chen
Orgul D. Ozturk

OCTOBER 2017

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11114

Occupational Match Quality and Gender over Two Cohorts

John T. Addison
University of South Carolina, Durham University Business School, and IZA

Liwen Chen
University of South Carolina

Orgul D. Ozturk
University of South Carolina

OCTOBER 2017

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Occupational Match Quality and Gender over Two Cohorts

Job mobility, especially early in a career, is an important source of wage growth. This effect is typically attributed to heterogeneity in the quality of employee-employer matches, with individuals learning of their abilities and discovering the tasks at which they are most productive through job search. That is, job mobility enables better matches, and individuals move to better their labor market prospects and settle once they find a satisfactory match. In this paper, we show that there are gender differences in match quality and changes in match quality over the course of careers. In particular, we find that females are mismatched more than males. This is true even for females with the best early-career matches. However, the direction of the gender effect differs significantly by education. Only females among the college educated are more mismatched and are more likely to be over-qualified then their male counterparts. These results are seemingly driven by life events, such as child birth. For their part, college-educated males of the younger cohort are worse off in terms of match quality compared to the older cohort, while the new generation of women is doing better on average.

JEL Classification: J3, J16, J22, J24, J31, J33, N3

Keywords: multidimensional skills, occupational mismatch, match quality, wages, gender wage gap, fertility, fertility timing

Corresponding author:
John T. Addison
Department of Economics
Darla Moore School of Business
University of South Carolina
1014 Greene Street
Columbia, SC 29208
USA

E-mail: ecceaddi@moore.sc.edu

I. Introduction

It has been shown that job mobility, especially early in a career, is the main source of wage growth. There is an influential literature that attributes this outcome to heterogeneity in the quality of employee-employer matches and to how individuals, through mobility, learn about their abilities and discover the job in which they can be most productive. As individuals are less mobile later in their career, these early-career moves can have lasting wage effects. We add to the literature that formulates match quality using data on occupational tasks and skill requirements and the pre-market skills of workers. This approach captures occupations as combinations of tasks and skills instead of discrete categories which then enables us to qualify and quantify (i.e. capture the direction and magnitude of) the change in quality of the match when a worker moves and the corresponding change in earnings.

Figure 1: *Early Career Match Quality Has Lasting Effects*[1]

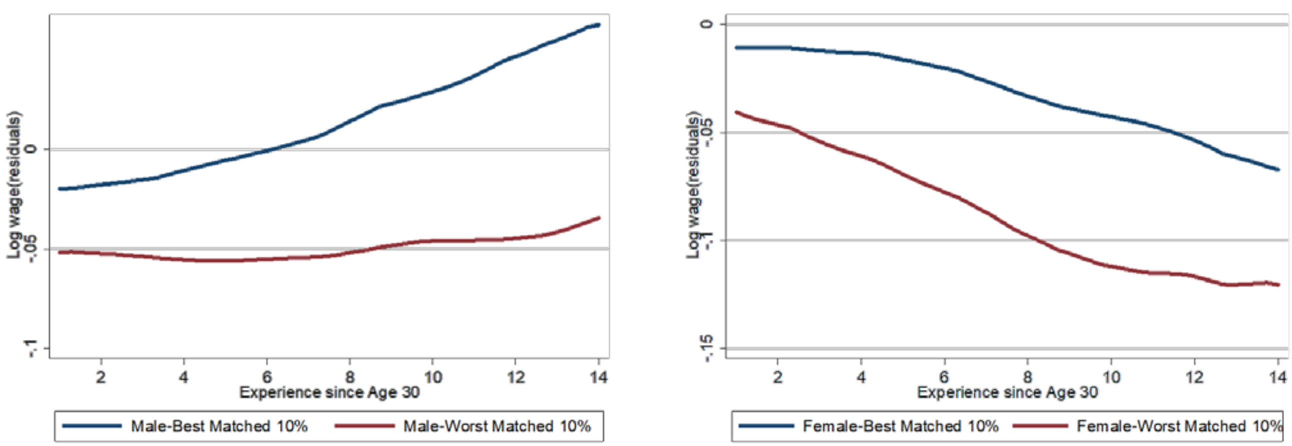


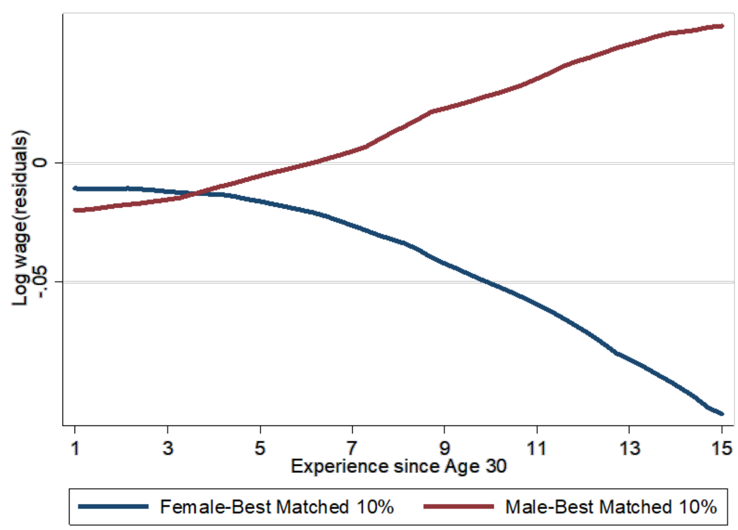
Figure 1 gives a preview of our motivating result from basic wage regressions supportive of this narrative. We rank individuals in the NLSY79 cohort core sample according to their average match quality in the jobs they held until age 30 – call this early-career match quality – and compare the residuals from log wage regressions for the best and worst matched groups (taking males and

[1]We grouped workers in terms of their average match quality up to age 30. The "best matched 10%" is the group of workers who had on average the lowest amount of mismatch (namely, the bottom decile of the mismatch distribution) and the "worst matched 10%" those who had the highest mismatch (top decile). Residual wages are obtained using the regressions reported in Table 11. Information on variables and definition of match measures is given in Section III. To construct the two sets of lines shown in the figure, we ran local polynomial regressions with residual wages on labor market experience for each group of workers, with a rule-of-thumb bandwidth.

females separately) for the balance of their labor market experience. The consequences of these differences in match quality are lasting for both genders. This result for males has been previously reported (e.g. Guvenen et al. 2016), but we now show that there is a permanent early-career match effect for females as well.

Finding the right skill match may require trying out different occupations. This implies search and turnover, both of which are costly. The process may be more onerous and costlier for workers who are constrained in their search parameters and flexibility to move. Traditionally, women have been expected to do more of the housework and care for dependents. Moreover, they have historically assumed a secondary role in family allocation decisions and mostly 'followed' their spouses. As a result, unlike their male counterparts, women may secure worse matches when they move. We would therefore anticipate greater mismatch in the labor market for women, with associated productivity losses even late into their careers. This may explain some of the gender wage disparities. Figure 2 illustrates the point by combining results for best matched males and females. Even among those workers who were successful in achieving good job matches early in their careers, we can see an unexplained wage gap that is increasing in experience.

Figure 2: ***Women Do Not Maintain Early Career Good Matches***



Our main contribution resides in the analysis of gender differences in match quality over a career. However, we also add to the literature by bringing the cohort dimension to the analysis of match quality. Female labor market participation has increased significantly over the last several decades. Moreover, the labor market is progressively less segregated by gender as women increasingly penetrate once male-dominated occupations. These developments are driven not only by technological advances that now make it possible for women to perform many 'physical jobs' without the exertion of physical power but also through their increased educational attainments. And even though college majors remain highly segregated by gender, there is undoubtedly a greater female presence in the technical and professional fields. Moreover, there has occurred a shift in perceived gender roles and in the formation and nature of relationships. Although less strongly encouraged and enabled than in Europe, women in the United States are expected to do it all and lean in, especially the more highly educated among them. In this paper, we shall also seek to determine whether such changes are reflected in labor market outcomes for women. These suggested shifts in the social environment will be addressed by using both the NLSY79 and the NLSY97. If younger cohorts face a more balanced division of labor, they may be better matched than their counterparts in the NLSY79. We might also expect the current generation of females to gain more from good matches if their careers are subject to less interruption. To the best of our knowledge, the present paper is the first to provide comparisons across cohorts.

To review our findings, we report that college-educated females have the highest amount of mismatch and those with the best early matches among them experience the greatest disparity in wages compared to their male counterparts. We also find that these results are driven by life events, such as child birth. Thus, females who have given birth do worse than (a) their male counterparts with a similar life event, (b) their female counterparts with no such life changes, and (c) compared to themselves before experiencing this birth event. For its part, delayed fertility reduces the negative effects of child birth on match quality. We show that there is a positive relationship between need for flexibility and mismatch, providing support for the personnel economics view of mismatch à la Goldin (2014). We also find that the younger cohort of females is experiencing less mismatch, possibly sharing the responsibilities for raising children more with their partners as the males in this cohort are doing worse. Past match quality is at least as important as current mismatch for its negative wage effects and the highest gender disparities in mismatch are in mathematical and technical skills which carry be the strongest wage implications.

The plan of the paper is as follows. Section II briefly section reviews the literature. Section III outlines the data sets used in the analysis and describes our manner of sample selection and variable construction. In constructing a theoretical framework and offering an introductory look at the data, Section IV provides motivation for the paper. It is followed in Section V by a detailed presentation of our empirical analysis. Section VI concludes.

II. Literature Review

Perhaps the best starting point in discussing the literature on skill mismatch is Goldin (2014) as she squarely confronts the gender component, addresses the issue of persistence, and provides evidence on nonlinear pay and the gender gap in earnings. Goldin is concerned to answer why the convergence in earnings between the sexes has still left an unexplained portion that cannot easily be laid at the door of discrimination, limited ability to bargain, differential employer promotion standards given gender differences in the probability of quits, or a lesser desire to compete on the part of females.[2] Something else is involved, as is hinted at by the fact that differences in hours worked and different amounts of time spent outside the labor force have dissimilar impacts on the time-adjusted earnings of different occupations. For Goldin (2014: 1094), as human capital endowments of the genders have increasingly converged, "what remains is largely how firms reward individuals who differ in their desire for various amenities." She argues that jobs for which bargaining and competing matter most have the greatest nonlinearities in pay with respect to time worked. Men and women begin their employment with rather similar earnings (women's hours adjusted earnings are roughly 90 percent of those of men) but this happy relativity does not last and the difference in earnings by gender increases markedly in favor of males during the first decades of working life.

For its part, gender mix is of secondary importance as the majority of the earnings gap stems from within- rather than between-occupation differences. In her empirical analysis, Goldin examines gender differences in pay for the 95 highest male income occupations (grouped into 5 categories [Business, Health, Science, Technology and 'Other'] using the American Community Survey, and charts residual gender differences by occupation. Business occupations have largest negative gender coefficients, Technology and Science the smallest. The trail thus returns to ascertaining the features of occupations that have high and low residual differences by gender, having addressed potential selection biases in Technology and Science.

---

[2] For the extensive literature on gender differences in competitiveness see, inter al., Niederle and Vesterlund (2007, 2010), Gneezy and List (2013), and the very recent review of Böheim, Grübl, and Lackner (2017).

Goldin offers what she terms a *personnel economics view* of occupational pay differences.[3] Some workers desire the amenity of flexibility or lower hours and some firms may find it cheaper to provide that flexibility. Thus, individuals place different values on the amenity of temporal flexibility while firms or sectors confront different costs of supplying that amenity. As a result, the hours-wage relation may be nonlinear and convex. It will only be linear when there are perfect substitutes for a particular worker – and zero transaction costs are involved. As an example of the latter, where there is no earnings premium for the number or timing of hours, she cites pharmacy jobs where earnings are found to be almost linearly related to time worked.

In common with much of the literature, Goldin explores the relation between wages (strictly, the residual gender gap) and occupational features using detailed job descriptions from the Occupational Information Network (O*NET). Five of the O*NET characteristics are singled out as most relevant for the model. These reflect time pressure, the need for workers to be around at particular times, the flexibility of the occupation with respect to scheduling, the groups and workers with whom the employee must regularly keep in touch, and the degree to which the worker has close substitutes. Higher values of the characteristics indicate greater costs of the amenities of reduced hours and more flexible employment. Scatter plots of the simple mean of the O*NET characteristics for each of the 95 occupations against the mean adjusted earnings gap for each occupation evince a negative association; that is, the female-to-male earnings ratio is lower, the higher the value of the characteristics. The relation is further pursued by examining evidence on the widening gender gap with age and differences in the gender gap by occupation using data sets specific to occupations and degrees. The occupations (degrees) are *Business* (MBAs), *Law* (JDs) and *Pharmacy*. Data for the two former occupations offer evidence of large increases in gender pay gaps with time since graduation and the desire for time flexibility as manifested in the arrival of children. Lower hours lead to lower earnings in a nonlinear relation. Lower earnings in turn yield lower participation rates, especially among those with higher-earning spouses. Pharmacists on the other hand have pay that is more linear with respect to hours of work, while those pharmacists with children often work part time and remain in the labor force. This is also typical of other occupations. Examples of occupations and sectors that have moved in the direction of less costly flexibility include healthcare, retail sales, banking, brokerage, and real estate. Goldin

[3] That is, differences in pay arise because of productivity differences at the workplace rather than by reason of inherent differences in human capital across workers.

concludes by identifying factors that have made pharmacists better substitutes for one another, noting that the potential for other occupations is considerable.

The literature on mismatch has painted with a narrower brush than Goldin and in the process either fails to deal with gender or to recognize the issue of persistence. In what follows, we shall primarily examine U.S. studies using the NLSY and O*NET but preface this material with some remarks on other studies. The preponderance of mismatch studies has looked at matching from the perspective of a discrepancy between a worker's highest completed schooling and the years of education required for a job (for an overview of this research, see Hartog 2000), the bases for which can be found in human capital theory and search theory with imperfect information, assignment models, as well as neoclassical models of household specialization. One issue with this 'over-education' literature is that it fails to consider both educational and skill mismatch together, leading to potential omitted variables problems. Another issue is that this literature has almost uniformly reported an absence of substantial gender differences (see Battu et al. 2000). Studies making a break with the over-education literature (i.e. offering alternatives to a match based on years of schooling and the schooling required for a job) include Robst (2007), Mavromas et al. (2013), Johansson and Katz (2007), and Liu et al. (2016).

Robst (2007) considers mismatch from the viewpoint of whether *the field of study at college* is related to the current job. Respondents reporting that their accepted job is not related to their major are adjudged mismatched. Having defined mismatch in this way, Robst next considers gender differences in worker-reported reasons for accepting a job outside the field of study and links such differences to wage disparities between mismatched men and women. Respondent reasons for accepting a job are deemed either demand- or supply-related. The former is captured by the response "job in highest degree field not available" while the latter comprise the categories "pay and promotion, career interests, working conditions, job location, and family-related reasons" that are subsequently grouped into *career-oriented* (the first two reasons) and *family-oriented* (the last three) which may therefore represent both individual job preferences and constraints on job search. Using data from the 1993 National Survey of College Graduates, results are provided for extent of mismatch, reasons for mismatch, and consequences of mismatch by gender. Although gender differences in the extent of mismatch are modest – 19 percent of men and 21 percent of women report that their work and field of study are not related – there are material differences in the reasons for mismatch, with women more likely to report amenity/constraint reasons for mismatch (principally family reasons) and men tending to report career-related reasons. There are modest though nonetheless statistically significant differences in the wage consequences of mismatch by gender (8.9 percent for females and 10.2 percent for males). That said, wage losses by reason for mismatch are larger, ranging from 17 to 21 percent for females and 18 to 29 percent for men.[4] Limitations of this study include problems associated with self-reporting, and the endogeneity of the relation between the reasons for accepting a position and wages.

A more thorough-going attempt to distinguish between the components of job mismatch is the study of Australian graduates by Mavromas et al. (2013) who distinguish between education and skill mismatch within a panel estimation framework. They consider the relation between over-skilling (though not under-skilling) and over-education (albeit this time derived using the *empirical method* where the education of an individual is compared with either the mean or modal level of education in the relevant occupation). In an interesting aside on the reason why workers may be mismatched in the case of education, the authors note that low ability individuals for their level of education may be over-educated but not necessarily over-skilled; that individuals may choose a job in which they are over-qualified because it offers them compensating advantages (e.g. less stress); and that employers may prefer over-educated workers because they are more productive and learn quickly. Note that in the latter case such individuals may suffer no pay penalty and their mismatch may be temporary if they are promoted relatively quickly. As for skills mismatch, it may accompany slack markets, or an environment in which firms lack well-developed hiring practices. The authors use longitudinal data from the first seven waves of the Household Income and Labor Dynamics in Australia (HILDA) survey, for employees holding a university degree or equivalent. Over-skilling is measured on the basis of the survey's 7-point scale giving the perceived degree to which the respondent feels that he/she uses many of his/her skills and abilities in the current job. As noted earlier, over-education is derived using the empirical method. In addition, the authors use a question in HILDA denoting how satisfied individuals are with different aspects of their job, this time on a 10-point scale. Mismatched workers consist of those who are over-educated, over-skilled, or both. It is reported that for women there is a wage penalty for each category of mismatch (relative to the well matched) while for males a wage penalty attaches only to those who are over-educated *and* over-skilled. Job satisfaction is unrelated to over-

[4] Note that they are positive for both genders in the case of mismatching due to pay and promotion opportunities.

education but is reduced by over-skilling and by over-skilling in combination with over-education. It is thus concluded that for many individuals, over-education is a question of choice or necessity whereas over-skilling is involuntary. The use of panel estimation methods strongly reduces the wage effects of mismatch observed in cross section. Finally, the wage penalty of mismatch is higher for females as it their reported dissatisfaction from the mismatch especially in the case of over-skilling.

The paper by Johansson and Katz (2007) examines skill mismatch and its impact on gender differences in the wage gap and the returns to education in Sweden, 1993-2002. The paper is in the spirit of ORU (Over-/Required-/Under-education) models, namely those drawing a distinction between an individual's attained level of education and the education required for his/her occupation. However, the authors use a job analysis measure of 'skills' mismatch rather than self-reporting methods or the use of average education or modal education, while noting its limitations (i.e. job evaluations are infrequently updated and may be broader than the actual work involved). Johansson and Katz estimate ORU-augmented Mincerian hourly wage equations for males and females separately, using each round of the Swedish Household Income Survey (HEK), 1993-2002, and seek to break down the gender gap in log wages (and changes in the gender gap) into its component parts using Oaxaca-Blinder (Juhn-Murphy-Pierce) decomposition(s). Skill mismatch is based on the normal educational requirements for the occupation according to the Swedish socio-economic index, and levels of education are defined according to the then applicable 5-element Swedish SUN classification. Appropriate or adequate education obtains if the two codes match; where they do not imputation is used to set the usual number of years to which the two levels of education correspond and thence by how many years the individual's education is adrift. On average, about half the sample has the required level of education, while 19 (31) percent is under-(over-) educated.

Focusing on the findings with respect to schooling and over- and under-education,[5] the authors report that the returns to education over and above what is required for the occupation are positive but smaller than the returns to required education (implying a reduction in the returns to schooling for each year of over-education) while those to under-education are again positive but below those for actual years of education (implying the additional premium the under-educated

[5] We do not discuss the JMP decomposition exercise because of a largely unchanged overall gender wage gap.

receive for each year of education that is normally required above that attained). Women are more likely to have more formal education than is required and conversely in the cases of men. The decomposition exercise shows that skill mismatch contributes to the increase in the gender gap because the decreasing effect on the gap of their longer schooling is in practice offset by the twin facts that their returns to jobs requiring more education are lower than for men and because women get less qualified jobs than do men with the same length of schooling. Skill mismatch accounts for considerably larger shares of the endowment term than traditional human capital variables and is between one-half and one-third of the industry (=segregation) effect.

A final Scandinavian study by Liu at al. (2016), using Norwegian data on college graduates, 1986-2007, is notable less for its particular adjusted earnings-based definition of skill mismatch (reflecting the mismatch between the (heterogeneous) skills supplied by college graduates and the skills demanded by (heterogeneous) hiring industries) than for its attempt to determine how mismatch varies over the business cycle. (On the separate literature showing that labor market conditions upon entry have large and persistent effects on careers, see, for example, Kahn 2010). It is reported that there is a strong countercyclical pattern to skill mismatch. Among graduates entering the labor market during a recession there is an increased probability of a mismatch and a decline in the average quality of job matches. The impact of initial market conditions on mismatch declines over time but persists over early careers and more so for poorer (lower IQ) graduates and those majors with cyclical demands. Initially mismatched graduates, defined as those who are matched to the wrong industry in their first job, suffer persistently worse labor market outcomes as a result of recession than their counterparts who graduate in good times. The distinct policy recommendations of this study focus on encouraging industry and job mobility among cyclically mismatched workers as a means of recovering initial losses.

We next turn in conclusion to two U.S. studies using the same basic combination of NLSY79 data and O*NET data as the present study.[6] The skill requirements of occupations

[6] A third study using the NLSY79 and addressing the effects of match quality on wages (and career decisions) is by Yamaguchi (2010), who does not measure job quality directly but instead infers it from the type of worker mobility. Workers are envisaged as searching for a better career match as well as a better employer match. Match quality changes are equated with a change in both employers and industry, while career match changes depend on a change in employers alone. The paper estimates a structural model of career decisions in which workers search for both careers and firms that are a good match for their idiosyncratic skills. The distinction drawn is between high school graduates and college graduates. The former group is more likely to search for jobs across careers than the latter, leading the author to conjecture that college graduates learn about their careers before they enter a labor market unlike high school graduates who learn about careers by changing jobs.

observed in the NLSY79 are obtained from O*NET database while initial worker skill bundles are obtained using ASVAB scores. In the first study, Lise and Postel-Vinay (2016) exploit the notion that workers are endowed with bundles of skills that are used in different proportions according to the task that they perform. The authors expand the standard search model of individual careers – mismatch in the model arising from search frictions – to allow for multidimensional skills and on-the-job learning. The model is used to throw light on the origins and costs of mismatch along 3 skill dimensions: cognitive, manual, and interpersonal. These skills are very different productive attributes with different returns and adjust differently, this heterogeneity cautioning against the use of a single scalar index of worker productivity. The costs of mismatch are very high for cognitive skills and are asymmetric; that is, employing a worker who is under-qualified in cognitive skills is said to be much costlier than employing an over-qualified worker. In assessing the production/wage cost of skill mismatch, the focus of this study is upon differences between skill categories in speed of human capital appreciation and decay.

The authors' structural model of job search with multi-dimensional job and worker attributes is estimated by indirect inference. In fitting the model, it is reported that jobs requiring higher levels of cognitive skills also tend to require high levels of skills in at least one of the other two dimensions, principally interpersonal. As far as adjustment is concerned, cognitive skills are not easily accumulated or lost, with a half-life of 7.6 years to learn and 28 years to forget. Manual skills adjust much faster, taking 16 months to acquire and 16-24 months to forfeit. For their part, interpersonal skills can be treated as fixed worker traits. Manual skills have relatively low returns, cognitive skills much higher returns, with interpersonal skills somewhere in between. Skill mismatch is most costly in the cognitive dimension and directionally where the worker possesses lower skills than required by the job.

The distribution of skills as the cohort of workers accumulates experience favors cognitive skills acquisition, whereas on average manual skills are lost. This pattern reflects the fact that jobs with high cognitive skills are intrinsically more productive, even if, in the process, workers are materially over-skilled in the manual dimension. Examination of the joint distribution of skill bundles and skill requirements points to positive sorting, the strength of which increases as workers accumulate experience. A corollary is that skill mismatch is stronger in the manual than in the cognitive dimension. That said, a substantial number of workers are found to be under-matched/over-skilled in the cognitive dimension. Finally, what social gain might be realized if workers could initially be placed in their preferred job? Lise and Postel-Vinay's counterfactual experiment puts the costs of the labor market frictions that create this mismatch at 8 to 22 percent of career output, depending on initial worker skills. The cost is increasing in manual and interpersonal skills though not monotonically with initial cognitive skills.

The final study considered here by Guvenen et al. (2016) most resembles our own.[7] The authors seek to evaluate the impact of skill mismatch on wages and patterns of occupational mobility. Familiarly, skill mismatch is based on the discrepancy between the portfolio of skills required by an occupation and the portfolio of abilities possessed by a worker for learning those skills, and as before worker abilities are obtained from the ASVAB and occupational requirements from O*NET and the occupational categories of the NLSY79. The authors derive a summary measure of mismatch and three dimensions of mismatch, namely its verbal, math, and social components (on the mechanics of the various aggregations, see section III). Mismatch in this model arises not from search frictions, as in Lise and Postel-Vinay (2016), but rather from workers having imperfect knowledge about their own skills, sorting into occupations that they consider optimal from the perspective of their perceived skill bundle as opposed to their true skill bundle. Albeit subject to subsequent updating, workers may overestimate (underestimate) their ability to learn a particular skill, causing them to choose an occupation with skill requirements for that type of skill that are too high (low) relative to their true ability. Empirically, skill mismatch is used as a regressor in human capital wage regressions and in statistical models of occupational switching. First, it enters wage regressions because mismatch is supposed to detract from human capital accumulation and hence depress the level of and rate of change in wages; in addition to which current wages will also suffer from past mismatch in previous occupations. Second, it is deployed in modeling occupational switching because the probability of changing occupations is increasing in mismatch as each wage observation engenders a bigger update in worker beliefs when mismatch is elevated. The ordering of occupational switches is also considered.

In documenting the empirical findings of the study, we focus on those for the summary mismatch measure. A baseline augmented Mincerian wage equation containing the mismatch

[7] However, for an interesting treatment using Canadian data in conjunction with O*NET data that considers the effect of economic downturns on over-qualification, see Sommerfield (2015). The author links heightened mismatch of this type to jobs formed in recession having relatively more manual tasks.

variable is first estimated. It is next supplemented with a mismatch interaction with occupation tenure and, in a final iteration, with a cumulative mismatch argument. (All estimates instrument for experience and tenure.) Mismatch alone has a strongly negative effect on wages. The tenure effect is negative indicating that mismatch not only lowers initial wages but also leads to reduced earnings growth over the duration of the match. For its part, cumulative mismatch has a negative effect on wages that actually serves to displace the effect of current mismatch. It is also shown that wage losses are sharper the larger mismatch. Specifically, the difference between the 90th percentile and the 10 percentile of mismatch is approximately 4.4% after 5 years of occupational tenure, widening to 7.4% after 10 years. Cumulative mismatch yields a wage difference of 8.9%. Being over-qualified mostly slows wage growth rather than having an immediate effect on levels. On net, the wage level analysis supports the notion that match quality affects the returns to tenure. A secondary result of interest is that the negative effects of mismatch vary by education, being much larger for college graduates, especially for the cumulative mismatch measure which roughly doubles in absolute magnitude. As far as occupational switching is concerned, baseline estimates controlling for the potential endogeneity of occupational tenure indicate that the effect of current mismatch on the probability of switching is positive and significant at the 0.01 level. Further, a worker who is at the 90th percentile of the mismatch distribution is 3.4 % more likely to switch than his counterpart in the 10th percentile. This value amounts to one-fifth of the average switching rate.[8] Finally, with respect to *switch direction,* it is reported that workers whose abilities exceed the skill requirements of their occupations, the over-qualified, tend to switch to occupations with higher skill requirements, and conversely for under-qualified workers. In addition, switches tend to correct past mismatch. That is, the more mismatched the worker in in the previous occupation the bigger the difference between the skill requirements of the last occupation and that of the current one. As a practical matter, workers who are over-qualified increase the skill requirements in the next occupation by less than the amount by which underqualified workers reduce them.

The modern literature based on a clearer distinction between the portfolio of skills required by an occupation and the portfolio of abilities possessed by the worker offers an important technical advance on approaches allied more closely with the over-education literature. On the other hand, an important lacuna of this new body of research is its narrower focus and failure to

[8] Differentiating by source of mismatch, the corresponding difference in mismatch is about 2% for verbal and math skills but to all intents and purposes zero for social skills.

consider women – an omission justified in the interests of limiting heterogeneity (sic). The present treatment seeks to re-insert gender into the contemporary methodology.

III. Data, Sample Construction, and Measurement Issues

(a) Data Sources and Sample Construction

Our main datasets consist of the 1979 and 1997 cohorts of the National Longitudinal Survey of Youth, namely the NLSY79 and the NLSY97. The former provides a nationally representative panel of data for the cohort of individuals aged 14 to 22 years in 1979, and the latter for youths aged 12 to16 years as of December 1996. Both cohorts were initially interviewed annually – the NLSY79 until 1994 and the NLSY97 until 2011 – but are now followed biennially. We restrict our sample to the core samples of both surveys, and thereby exclude the military as well as the oversample of Hispanic, black, and low-income youth. We further restrict our sample to include only those individuals who work in dependent employment and who are not self-employed or working for non-profit organizations. We also exclude those who work for no pay or who report hourly wages of less than $1. Having also excised those with missing information on any of the variables used in the analysis, as well as observations for which the wage entries are clearly in error,[9] our final sample comprises 42,022 person-year observations (1890 males and 1980 females) from the NLSY79 and 15,893 person-year observations (from 1588 males and 1434 females) from the NLSY97 over the survey periods analyzed. Table 1 reports the observation losses due to each sample inclusion criterion.

[Table 1 near here]

In addition to its long panel nature, use of the NLSY has two other advantages. The first is that it effectively tracks workers' actual labor market experience, allowing us to correct for any measurement error in the conventional imputed measure based on age and education (i.e. age – schooling – 5). The second is that it allows us to control for ability (and skills of the individuals across several dimensions), using the Armed Services Vocational Aptitude Battery (ASVAB) test scores, which measures are unavailable in other panel data sets of a similar nature. We use the age-adjusted percentile scores of respondents on the subtests of ASVAB as basis of our individual skill measures (see the next subsection).

[9]For example, we have a few instances of wage growth of more than 100%, followed by huge declines in the next period yet unaccompanied by any material change in job characteristics.

Although labor market activity has been recorded in great detail in both surveys since their inception, the occupations and industries are not coded consistently across each wave of either survey. Until 1981, all occupations and industries in the NLSY79 were coded using 1970 Census codes (Census Occupational Classifications/COCs and Census Industrial Classifications/CICs, respectively). Beginning with the 1982 survey, occupations were coded using the 1980 codes, in addition to the 1970 codes, until 2002. After that year, the 2002 COC was used to code occupations,[10] and after the 2010 round the 2010 COCs were also provided. For its part, even though the first five rounds of NLSY97 employed 1990 codes for occupation classification, 2002 Census codes were added retroactively for all rounds and are also provided for the newer rounds of the survey along with 2010 COCs. Similarly, industries are described by their 3-digit 1980 CIC in the NLSY79 until 2000 and thence 4-digit 2002 CICs are used. 2002 CICs are available for all rounds of the NLSY97. We mapped all available NLSY79 and NLSY97 occupation codes to be able to study the full extent of the data panel available.[11]

Table 2 provides basic descriptive statistics for our NLSY samples.

[Table 2 near here]

(b) Measuring Match Quality

Determination of Worker Skill Endowments and Occupational Skill Requirements

We define individual workers' skill mismatch as the discrepancy between their premarket skill levels and the requirements of the occupations in which they are employed. In linking the skill supply side (viz. workers' endowments) with the demand side (occupational requirements), we exploit the tools developed by the ASVAB (Armed Services Vocational Aptitude Battery) Career Exploration Program. The ASVAB Career Exploration Program is administered by the Department of Defense (DoD) with a view to helping ASVAB participants identify and explore suitable career possibilities in the private, public, or military sectors. Both NLSY surveys administered the ASVAB tests around their inception; specifically, for the first round of NLSY97 and the second year of the NLSY79. All NLSY79 respondents and about 80 percent of the

[10] The occupational classification released and used in Census 2000 was slightly revised by the Census Bureau in 2002. But these two sets of codes, the 2000 COC and 2002 COC, are essentially the same. In the NLSY datasets, 2000 codes were used in 2002, but for the 2004 rounds of NLSY surveys the 2002 version of the coding system was utilized.

[11] See Appendix 1 for more details on occupational code mapping.

NLSY97 sample participated in the computer adaptive test of the Armed Services Vocational Aptitude Battery (CAT-ASVAB).[12]

We consider four categories of skills (or Knowledge/Skills/Abilities, or KSAs): Mathematical, Verbal, Science/Technological/Mechanical (STM) and Social. For the first three categories, for all those in the NLSY samples with valid test scores, we constructed composite measures using percentile ranks on select ASVAB subtests. Specifically, for verbal skills we used the percentile scores on *Word Knowledge* and *Paragraph Comprehension*, for mathematical skills the scores on *Arithmetic Reasoning* and *Mathematical Knowledge*, and for STM skills the scores on *General Science*, *Mechanical Comprehension*, and *Electronics Information*.[13] Next, using the weights provided by the NLS,[14] we created a comparable composite skills measure from these subtest scores for each NLSY respondent. We then standardized these skill percentile ranks to be between 0 and 1.

For the construction of the remaining endowment measure – social skills – we follow a strategy that combines the methods used by Deming (2017a) and Guvenen et al. (2016). We use two questions from the NLSY79 survey (specifically, the third round of the survey in 1981) where respondents are asked to report on their then current sociability and sociability at age 6 (retrospective) along with their rank on the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale.[15,16] The NLSY97 does not ask these sociability questions nor does it collect data on

[12] For details of the administration of the ASVAB and CAT-ASVAB tests, the reader is referred to the NLSY79 and NLSY97 web pages: respectively, https://www.nlsinfo.org/content/cohorts/nlsy79/topical-guide/education/aptitude-achievement-intelligence-scores and *https://www.nlsinfo.org/content/cohorts/nlsy97/topical-guide/education/administration-cat-asvab-0*.

[13] This approach is similar to that used by Guvenen et al. (2016) other than for the inclusion of STM scores. There is no consensus in the literature on construction of the ability measures, and even though almost all studies utilize ASVAB test scores, they select different ability dimensions or different subtests for measurement of these dimensions. We have checked the robustness of our results to variation in measurements, such as the exclusion of STM skills by Guvenen et al. (2016) and the restriction of ASVAB measured abilities to *cognitive* and *manual* by Lise and Postel-Vinay (2016) who also analyze mismatch by separate ability dimensions as opposed the use of an aggregate measure.

[14] We thank Steve McClaskie and other NLS program staff for their help in this connection.

[15] The Rosenberg Self-Esteem Scale is a measure of self-worth while the Rotter-Locus of Control Scale is designed to measure the extent to which individuals believe they exercise control over their lives (the predominance of self-determination over chance or fate). For the NLSY79 cohort, tests of these two endowments were administered in 1979 and 1980, respectively. Similar to the other dimensions of KSAs, the literature displays multiple ways of measuring social skill or abilities. In Guvenen et al. (2016) the social skill endowment is measured using the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale. These authors refer to the measure as indicating social ability, whereas Deming (2017a) uses the label non-cognitive skills. Deming in fact uses the sociability questions for the NLSY79 cohort and extraversion measures for the NLSY97 cohort as his social skills measure. Our results are robust to alternative measures using either of these definitions.

[16] Deming (2017a) uses two additional questions on high school participation in clubs and sports for his analysis of 1979 cohort data. For his analysis of cohort differences, he switches to a two-question measure. We only use two sociability questions for the NLSY79 which is consistent with his cohort analysis.

the Rotter and Rottenberg Scales. Instead, respondents are asked a series of questions to determine personality traits (Big 5 Personality Factors). Following Deming (2017a), we use two questions on *extroversion* and two questions on *conscientiousness* to construct a social skill rank comparable to the NLSY79 cohort's measure. We downloaded the standardized measurements from Deming's (2017b) data file, and then converted the scores to percentile ranks for each cohort of NLSY respondents. Table 3 reports descriptive statistics on the skill endowments for each NLSY cohort by gender and educational attainment.

[Table 3 near here]

In our analysis, each and every occupation is defined by the combination of KSAs it requires. We use the O*NET database to determine the task requirements of each occupation.[17] For each of the ASVAB test scores used as components of the first three skill endowments, there is a corresponding occupational task which utilizes that knowledge, skill, or ability. The DoD has a mapping between ASVAB subtests and occupational tasks that describes how they arrive at their assignments. This mapping is provided in Appendix 2. However, there is no social skill component to these DoD assignments. Again following Guvenen et al. (2016) and Deming (2017a), therefore, we constructed the occupational requirements of social skills using the following descriptors "Social Perceptiveness", "Coordination", "Persuasion", "Negotiation", "Instructing", and "Service Orientation" taken from the O*NET database. We use the previously described occupational code mapping strategy for merging O*NET occupational characteristics to the NLSY data.

Mismatch

The extent of skill-mismatch is measured as the absolute value of the differences between the percentile-rank scores of an individual's skill endowments and the percentile-rank scores of skills required in that individual's occupation.[18] Specifically, let A_{ij} represent individual *i*'s percentile-rank-scores in the ASVAB test for skill j (where j denotes mathematical, verbal, scientific/technical/mechanical skills, and social skills). Recall that A_{ij} does not vary by year or an individual's occupation. Let R_{ijcy} denote individual *i*'s O*NET occupational requirements for

[17] We are using the 2007 version of the O*NET database, after Hirsh and Manzella (2015). We are indebted to Barry Hirsch for kindly providing us with these data.
[18] We also developed an alternative measure based on cosine similarity between vectors of skill endowments and skill requirements for robustness checks. Our results proved robust both to the use of this alternative measure as well as to measures using only three of the four KSAs (namely, math, verbal, and social), as in Guvenen et al. (2016).

skill j, in occupation c, in year y. The degree of skill mismatch for individual i for skill j, in occupation c, in year y is calculated as

$$q_{ijcy} = |A_{ij} - R_{ijcy}|,$$

such that the lower the value of q, the better the skills are matched.

[Table 4 near here]

Table 4 offers a descriptive view of the mismatch measures in aggregate and by skill type for each gender by cohort. The table contains several important broad indications. First, it is more likely that individuals have more skills than are required rather than less (compare columns [2] and [5]). Second, the magnitude of over-qualification is larger than the magnitude of under-qualification (compare columns [1] to [4]). Third, as a result, the probability of being significantly over-qualified (endowments more than one standard deviation above requirement) is more likely than being under-qualified (compare columns [3] and [6]). We also observe, as expected, that the severity of over-qualification (see column [1]) decreases over the course of individual careers for the older cohort.[19] Moreover, among workers with more than 10 years of experience, the share of those who are over-qualified is lower. On the other hand, the share of those who are under-qualified is mostly higher. In this paper, we mainly concentrate on the size of mismatch and effect of mismatch on wage outcomes. However, we will also create dummy indicators to indicate the direction of mismatch and test whether being over-qualified for a job is different from being under-qualified both in terms of its determinants and effects on wage outcomes.[20]

IV. Conceptual Framework

The conceptual framework guiding our thinking is similar to that outlined in Guvenen et al. (2016), but containing additional layers to accommodate gender and cohort differences. In their set-up, an individual worker's productivity is a positive function of match quality. As a result, all else equal, individuals will choose the job where they believe they are better matched. They update their beliefs about their abilities given the matches they experience and move to improve match quality, if and when they can.

[19] It will be recalled that the 1997 Cohort is still young, the average workers having less than 10 years of experience.
[20] We will define being over-matched as being *significantly* over-qualified; specifically, having an endowment level that is more than one standard deviation above the occupational requirements.

The first additional layer is inspired by Goldin (2014), who identifies differences in need for flexibility as an unresolved source of gender wage disparities. As noted earlier, some workers desire the amenity of flexibility or lower hours and some firms may find it cheaper to provide that flexibility. Thus, individuals place different values on the amenity of temporal flexibility while firms or sectors confront different costs of supplying that amenity. As a result, Goldin argues that the hours-wage relation may be nonlinear and convex. We would further argue that these flexible jobs are not offered across the range of skills and for all tasks. As a result, when life events such as birth of a child occur and alter preferences for flexibility, workers face a restrictive set of occupations that accommodate this need and they may end up in an occupation for which they are over- or under-qualified. In our view, those workers facing such flexibility/match quality tradeoffs are more likely to be women – and given our data patterns are more likely to be over-qualified. The tradeoffs result in wage losses not only by reason of compensating differentials, and the above-mentioned nonlinear and convex relationship between hours and wages, but also because the workers in question are underutilized in their jobs. Given the findings in the literature, they are more likely to be underutilized in terms of math and technical skills, which happen to be those with the highest wage rewards.

The second layer involves household decision making à la Frank (1978). If females are secondary breadwinners, once their husbands make optimal job search decisions involving a shift in location, the wife also moves regardless of the job opportunities at the new location: she is a 'tied mover.' Equally, the wife will be a 'tied stayer' if her husband has optimized his job search in the current location. In either case, the female partner may be expected to confront a worse match and a higher risk of being over-qualified. Even though we do not formally formulate and structurally estimate a model of flexibility and differential over-qualification, we test implications of the above framework using reduced form specifications.

V. Econometric Analysis

Our analysis proceed as follows. First, we document the determinants of the magnitude of mismatch in the labor market. For the older cohort we next seek to gauge the size of the gender difference and how life events – marriage, child birth, and their timing – affect worker-occupation match quality. We also test how the need for flexibility and household decision dynamics may play a role in this process. For the NLSY 79 cohort we then evaluate the cost of being mismatched in terms of lost wages, and ask how much of gender wage disparities by educational level can be explained by workers' history of match quality. We also look at the NLSY97 outcomes. Although this cohort is too young to have experienced extended labor market histories, we can compare it with the NLSY79 for early-career outcomes. Specifically, we test whether or not, all else equal, the two cohorts of the NLSY display different match quality by gender in this early-career phase. Finally, we analyze match quality and wage effects of match quality by skill type.

(a) Gender and Mismatch

In Table 5 the dependent variable is the standardized total amount of mismatch. The first column of the table includes only the female dummy. The second column includes some demographic characteristics and ability measures as well as year dummies. Column 3 contains results from a third specification with additional controls for occupation tenure and experience variables, as well as their interaction with ability measures, and will constitute our baseline equation for subsequent regressions. From the table we see that for the older cohort (NLSY79) females are more mismatched on average than their male counterparts, after controlling for a set of demographic and labor market-related variables. This pattern seems to be driven by the greater amount of mismatch among highly-educated females. Specifically, females with college degree or higher educational level are found to have 12 percent more skill mismatch (from the third column; 0.1184=0.1509 - 0.0325) compared to their male counterparts with the same educational level. For those individuals without a college degree, on the other hand, the gender difference in extent of mismatch (-0.0325) is not significant.

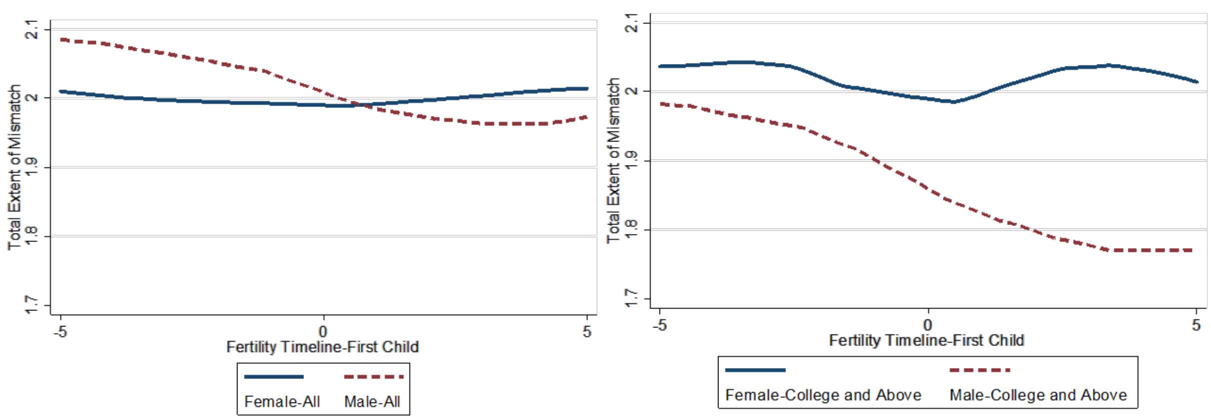
[Table 5 near here]

There is one obvious issue we need to consider here. Individuals with greater labor market attachment may have better match outcomes, while good matches may lead to longer tenures in occupations and more years in the labor market. We resolve this endogeneity problem by instrumenting for employer tenure, occupational tenure, and total experience. Following Altonji and Shakotko (1987) and Guvenen et al. (2016), we instrument individuals' employer tenure and occupational tenure with their relative position in the tenure hierarchy with a given employer or occupation, controlling for the possibility of multiple spells of employment with the same employer or in the same occupation.[21] Total labor market experience is instrumented in a similar fashion. Column (4) in Table 5 reports results for this IV specification: for the estimates that are of interest to us we see no significant change in magnitude or sign.

(b) The Role of Household Formation and Fertility

Whenever gender disparities are of concern, it is imperative to discuss how gender roles, family formation, and fertility contribute to the particular disparity in question. Although factually marriage does not seem to be correlated with mismatch, this is not the case for fertility. The descriptive association between fertility and match quality is graphed in Figure 3 for NLSY79 cohort. That is, normalizing the employment timeline of a worker around the birth of the first child, there is clear evidence of an up-tick in female workers' mismatch. Over the same timeline, male worker matches actually improve.

Figure 3. ***Fertility and Mismatch – First Birth Timeline***



Now this descriptive look at the fertility-mismatch relation is likely biased. Fertility incidence and timing are probably endogenously determined, while there is selection in who

[21] $IV_{emp} = T_{emp} - \overline{T_{emp}}$ and $IV_{occ} = T_{occ} - \overline{T_{occ}}$, where $\overline{T_{emp}}$ is the average duration for individual i with the same employer k and $\overline{T_{occ}}$ is the average duration for individual i with the same occupation j. $\overline{T_{emp}} = \frac{1}{T}\sum_{t=1}^{T} T_{emp\,i,k,t}$, where T is the total number of spells that an individual is observed with the same employer. $\overline{T_{occ}} = \frac{1}{N}\sum_{t=1}^{N} T_{occ\,i,j,t}$, where N is the total number of spells that an individual is observed with the same occupation j. Total experience is also instrumented in the same way, with an instrument $IV_{exp} = T_{exp} - \overline{T_{exp}}$, where $\overline{T_{exp}}$ is the average duration that individual i stays in the labor market $\overline{T_{exp}} = \frac{1}{S}\sum_{t=1}^{S} T_{exp\,i,j,t}$, where S is the total number of spells that an individual is observed to be in the labor market.

chooses to have a child. Controlling for possible confounding factors and addressing issue of bias, Table 6 further explores the effect of marriage and fertility on the size of mismatch. We run three sets of empirical models. The first set addresses the issue of endogeneity in respect of the tenure and experience variables (pooled IV).[20] The second set exploits the panel nature of our data and addresses the possibility of unobserved individual heterogeneity with a fixed effects (within variation) set-up. The last set combines efforts to alleviate both concerns and reports fixed effects-instrumental variables regressions.

[Table 6 near here]

The first column of pooled IV regressions in Table 6 suggests that females with at least one child are on average 10 percent of a standard deviation more mismatched than are males with at least one child. For females, having a child is associated with an increase in mismatch by about 7 percent (0.0664=0.0971-0.0307) of a standard deviation relative to childless females. The association with having a child is much smaller and in fact is not statistically significant for males (-0.0307). For its part, marriage is not significantly related to mismatch in the pooled IV regressions. However, when individual unobserved heterogeneity is factored in, both by itself and in addition to the endogeneity of tenure and experience variables, marriage is negatively related to mismatch for females. In the fixed effects specifications, the effect of fertility is more pronounced for women. Women who have at least one child are 10 to 11 percent of a standard deviation more mismatched than their childless female counterparts and 13 to 14 percent of a standard deviation more mismatched than their male counterparts.

[Table 7 near here]

Our descriptive statistics provided earlier in Table 4 suggested that mismatch mostly takes the form of being over-qualified for the occupation that one holds. In other words, the phenomenon of having more skills than can be utilized is more common in the data and arguably is the more common problem faced by females if family-related issues are the source. For this reason, the incidence of over-qualification is the outcome of interest in Table 7. Specifically, our concern is with “significant over-qualification,” which we define as having at least one standard deviation more in endowments than is required for the occupation. We report that the odds ratio of being over-qualified for females with children over females without children is 1.96. For males, on the other hand, having children does not significantly affect the odds of being over-qualified.

Moreover, marriage alone has no significant effect on the probability of being over-qualified for either gender.

[Table 8 near here]

Tables 6 and 7 show that main reason for mismatch among females is fertility, with children affecting female and male mismatch differently. To understand the dynamics of the fertility-mismatch relationship, in Table 8 we look at the change of mismatch along the fertility timeline for the NLSY79 cohort. For an interval extending up to six years after the first birth, females have worse matches than is the case for matches made within 3 years prior to that birth event. The disparity is approximately 10 percent of a standard deviation in the OLS results, 15 percent of a standard deviation the IV-FE regression, and up to 17 percent of a standard deviation in the case of the FE-only regression. The opposite holds for men, with the extent of mismatch declining significantly within the 6 years following the birth of the first child. More than 6 years after the birth of the first child, the extent of mismatch is even higher for females, but no significant changes are detected for men.

[Table 9 near here]

The above findings imply that the increment of mismatch shortly following the first birth may accumulate, leading to even greater mismatch later into a woman's career. It is therefore of interest to explore whether delaying fertility might reduce the size of mismatch brought upon by parenthood. To this end, we measure the "timing of the first birth" in two different ways: firstly, by the mother's age at the first birth; and, secondly, by the number of years that have elapsed between entry to the labor market and that first birth. Considering the possible endogeneity between fertility timing and quality of the occupational match, we now instrument for both timing measures using the age of the individual respondent's sibling at the first birth. In a second set of regressions (IV-2) we also instrumented for the individual's occupation and employer tenure and job market experience in the manner described earlier. The estimates are reported in Table 9. We see that delaying the time of birth significantly reduces the amount of mismatch for women (by about 3 to 4 percent of a standard deviation per year), but no significant effects are observed for men in either instrumental variables set-up. When comparing the last two sets of results with the OLS estimates, we can conclude that there is significant endogeneity between age at first birth and occupational mismatch. Specifically, individuals recording worse mismatch have earlier first births, or individuals with better matches delay fertility.

(c) Occupational Flexibility and Mismatch

In Table 10, we test whether a Goldin-type explanation has any purchase when it comes to mismatch. To this end, we constructed an occupational "flexibility score," which is the average of five O*NET working context measurements: *time pressure*, *contact with others, establishing and maintaining interpersonal relationship, structured vs. unstructured work,* and *freedom to make decisions.*[22] The higher the score, the more flexible is the occupation. We see from the first four columns of the table that the interaction term between the "have at least one child" dummy and occupational "flexibility score" is positive and significant for both males and female, where all estimations are based on the IV-FE model. This finding indicates that people with children tend to work in occupations offering high flexibility at the expense of a better skill match. Among individuals with at least one child, a one standard deviation more flexible job is associated with a (roughly) 4 percent of a standard deviation higher mismatch for males and a 6 percent of a standard deviation higher mismatch for females. As regards the timeline of birth, the last two columns of Table 10 indicate that for both sexes working in flexible occupations leads to more over-matching in the 6 years following the birth of the first child – by about 5 percent of a standard deviation in the case of males and about 8 percent of a standard deviation for females – relative to the up-to-three-year period before the first birth.

[Table 10 near here]

(d) Wages and Mismatch

Following Guvenen et al. (2016) we calculated the wage loss associated with mismatch, but for both genders rather than males alone. In Table 11 we see that the wage loss associated with the

[22] The five O*NET working context measurements are defined as follows: (i) Time pressure: how often does this job require the worker to meet strict deadlines? The higher the raw score, the lower the flexibility, (ii) Contact with others: how much does this job require the worker to be in contact with others (face-to-face, by telephone, or otherwise) in order to perform it? The higher the score, the lower the flexibility, (3) Establishing and maintaining interpersonal relationship: developing constructive and cooperative working relationships with others, and maintaining them over time. The higher the score, the lower the flexibility, (iv) Structured vs. unstructured work: to what extent is this job structured for the worker, rather than allowing the worker to determine tasks, priorities, and goals? The higher the score, the higher the flexibility, and (v) Freedom to make decisions: How much decision making freedom, without supervision, does the job offer? The higher the score, the higher the flexibility. Each element is standardized to have a mean 0 and a standard deviation of 1 in the O*NET data. To arrive at the flexibility score, we took the negative of the first three measurements, and obtained an average score for the five measure across all occ1990dd occupations. We then standardized this score to have a standard deviation of 1.

total extent of mismatch has three components: (i) a threshold penalty, (ii) a decreasing return along the career path, and (iii) a wage penalty associated with cumulative past mismatch.[23]

In the first set of regressions[24] we only include the measure of mismatch in the current occupation. Our results imply for males that someone who is matched one standard deviation worse (more mismatch) than the average mismatch amount vis-à-vis someone with a one standard deviation better match will earn 5 percent less (-0.0236*2 from the first regression and -0.0243*2 from the gender specific one). The wage effect is a little less for females at 4.4 percent ([-0.0236+0.0014] *2 from the first regression and -0.0227*2 from the gender specific one), even though the gender difference in the effect of mismatch on wages is not significant in this specification. When we add the occupation tenure interactions to capture differences in returns along the career path, the threshold effect declines but the change is not significant. However, when the cumulative mismatch – that is, the measured history of mismatch, as described above – is added to the model the effect of current mismatch becomes insignificant. Cumulative mismatch, on the other hand, implies a 12 percent (-0.0598*2) wage difference for males who have a one standard deviation better than average match history compared to those with a one standard deviation worse than average match history. Cumulative mismatch in this model does not have differential effects on female wages.

Next we add the education dimension by differentiating between individuals without a college degree and those with at least a college degree. Our results indicate that college graduates suffer from much stronger current and cumulative mismatch wage penalties than their less educated counterparts. One standard deviation higher than average current mismatch implies a wage penalty for college graduates of about 3 percent more on average and an even stronger cumulative mismatch effect of 5 percent more relative to a non-college graduate of the same level

---

[23] We calculate the cumulative mismatch in the same manner as do Guvenen et al. (2016); that is, as the weighted average of past mismatches where the weights are formulated as the length of tenure in a given occupation over the total labor market experience. Strictly speaking, this is not cumulative mismatch but rather "average past mismatch." However, to maintain consistency with the extant literature we shall continue to call it cumulative mismatch.

[24] Here we are only reporting the results of the instrumental variables specifications that we shall use later for our wage gap calculations. In these regressions, in order to distinguish the first year of the first spell with a new employer from the first years in later spells, an old job dummy is created (viz. $oldjob$ equals 1 if the current employer is an employer the worker had in the past, which will only be zero at the first year of the first spell) and also instrumented as are the occupation tenure and experience variables. Specifically, $IV_{oldjob} = oldjob - \overline{oldjob}$, where $\overline{oldjob} = \frac{1}{T}\sum_{t=1}^{T} oldjob_{i,k,t}$, where T is the total number of spells that an individual is observed with the same employer.

of mismatch. That is to say, a male college graduate who is mismatched consistently (in his current and previous occupations) one standard deviation worse than the average mismatch will earn 12 percent (-0.0421+-0.0311+-0.0513) less than a college graduate male who is of average mismatch, whereas the penalty is only about 4.5 percent in the case of a male non-college graduate relative to an average mismatched male non-college graduate. Once again, regardless of educational status, there are no significant gender differences in mismatch wage penalties.

[Table 11 near here]

Thus far, we have shown that the extent of total skill mismatch for females increases after the first birth and deteriorates when the first child is older and potentially there are more children in the household. We also demonstrated that cumulative mismatch has more detrimental effects than current mismatch. Although there are no significant gender differences in effect of mismatch on wages, greater current and cumulative past mismatch for females may contribute to the increasing wage disparities between males and females over a career. For this reason, in Table 12 we next link the gender difference in mismatch to the gender wage gap by creating career-long mismatch effects for the best and worst matched individuals in early career by college education and its absence. Early match quality is ranked according to an individual's mismatch over the first 5 years of experience. A distinction is drawn in the table between the top (worst matched) and bottom (best matched) deciles of the mismatch distribution. Current and past mismatch values are averages for these groups with the given level of experience and early career match quality. We learned from Table 11 for college graduates that the impact of both current mismatch and cumulative mismatch is significantly negative while wage growth is not significantly influenced by the degree of mismatch (captured by the coefficient of the occupation tenure-mismatch interaction). For our wage gap calculations, we first consider in the penultimate column of Table 12 only the precisely-estimated mismatch coefficients. The last column includes the imprecisely estimated coefficients as well, now including for college graduates the occupation tenure-mismatch interaction coefficients. According to our projections, for the college educated with the best early career matches their career record of mismatches is such that there is approximately a 7 percentage point difference in wages due to their current mismatch and mismatch history. This figure is much reduced for the less well matched even among college graduates (where it amounts to around 2 percentage points only) and is almost non-existent among the non-college educated.

[Table 12 near here]

(e) Cohort Differences

For this next part of our analysis we use only those NLSY79 observations for respondents aged are 33 years or younger. In this way, we have a sample comparable to the NLSY97, and can combine the two cohorts. We capture the differences across cohorts with a NLSY97 dummy both by itself and in interaction with other variables of interest. Table 13 examines cohort differences in mismatch and over-qualification. For the 'younger cohort' (NLSY97), we see that college-educated females continue to be more mismatched than the male counterparts, although not significantly so. On the other hand, among those without college degrees, females are less mismatched and less likely to be over-qualified than their male counterparts of both generations and also when compared to the older cohort of females. Finally, Table 13 suggests that the younger cohort of males perform worse, recording greater skill mismatch on average.

[Table 13 near here]

Table 14 compares the role of fertility and its timing on mismatch across cohorts. It is apparent that younger-cohort men display a greater amount of mismatch after having a child than does the older cohort of males, while the opposite is true for women of that cohort. Relative to the older cohort of males, they are more mismatched within 6 years after the birth of the first child when compared to the baseline interval before birth; however, this effect although large is not precisely estimated. This pattern is in strong contrast to what we have observed for older cohorts, wherein males become better matched after having a child, especially when children are young. This shift implies that male millenniums assume more family responsibilities than did the male baby-boomers. For younger-cohort females with at least one child, mismatch is less than that among their counterparts in the older-cohort. With respect to the fertility timeline, we observe that young-cohort females are also less mismatched more than 6 years after the first birth than the corresponding group of older-cohort females.

[Table 14 near here]

We next check to see if this slight shift in the gender burden of mismatch is due to the fact that women are today more engaged in the labor market, and have in the words of Goldin (2014) achieved a *grand convergence* in human capital attributes that has made them equally if not more productive than men. According to a 2017 BLS report, the share of women earning more than their husbands increased by more than 60 percent over the last several decades.[25] Table 15 seeks to determine whether this shift might underpin our observed cohort differences. It may also be seen as offering a test of Frank's (1978) theory of differential over-qualification. That is to say, in Table 15 we examine whether being a breadwinner decreases the level of mismatch both with and without children in the household. It is indeed apparent that not only are breadwinners less mismatched than non-breadwinners but also that the effect is stronger for NLSY97 cohort females, while weaker and not statistically significant for males of the younger cohort.

[Table 15 near here]

(g) Mismatch by Skill

As was noted in section II, the most recent literature suggests that the effects of mismatch might differ by the type of skill for which the worker is mismatched. In particular, wage penalties appear highest for cognitive skills. Moreover, cognitive mismatch wage effects seem to be more permanent. Workers can make up for the manual and social skills that they lack, but cognitive skills when under matched are hard to rectify. Also, when overmatched, unrealized returns to cognitive skills are higher than for all other skill dimensions. Thus, a pressing question in our own inquiry, is whether gender differences in mismatch are differentially greater for math skills; another is whether wage effects of mismatch are different as well.

Table 16 first examines gender differences in mismatch by skill type and education level, and how fertility might compound these gender differences. We see that among non-college graduates, females are about 5 percent of a standard deviation more mismatched than their non-college educated male counterparts when they have a child. They are also about 4 percent of a standard deviation more mismatched than their childless selves. Among college graduates with at least one child, females are about 10 percent of a standard deviation more mismatched than their male counterparts. There is an even larger gender difference in the amount of mismatch for STM skills in the presence of children; compared to their male counterparts who have at least one child, corresponding females are about 9 percent of a standard deviation more mismatched. This particular gender effect is even stronger in the case of college-educated individuals, although the

[25] https://www.bls.gov/opub/reports/womens-databook/2016/pdf/home.pdf; accessed on Oct 20th, 2017.

coefficient estimate is statistically insignificant. Women who have at least one child are also more mismatched in other skill dimensions compared to their childless counterparts.

The results for over-qualification, shown in the lower panel of the table, are frankly mixed. However, in the skill category of most importance – math (=cognitive) skills – college-educated women are also more likely to be overqualified when compared with both their male counterparts and childless women.

[Table 16 near here]

Finally, in Table 17, we seek to isolate the wage consequences of mismatch by skill type. As before, we look at the question both in terms of current and past match quality, and distinguish three sources of effect. Similar to our earlier results, when occupational tenure-mismatch interaction is considered, evidence of any threshold mismatch effect most disappears. Moreover, also as before, cumulative mismatch has the biggest wage penalties across all skill types. And similar to the findings of the literature, we report that mismatch in terms of math skills (or cognitive skills) has the largest wage effects for males. For females, on the other hand, penalties are lower for math mismatch but are of the same order of magnitude for cumulative mismatch on math, verbal, and social skills.

[Table 17 near here]

VI. Conclusions

We find that a significant portion of gender wage disparities – from around 1 percentage point for the worst matched at the beginning of a career, up to about 7 percentage points among individuals with best early matches – can be explained by match quality differences, especially among college-educated individuals. Mismatch may be driven by gender segregation and discrimination in certain fields given that the biggest mismatch gender differences are in mathematical and technical skills.

Another explanation for this disproportionate female mismatch is the division of labor/specialization in the household and traditional gender roles – and expectations in general. Our results indicate that, after giving birth, highly-educated women trade off flexibility for match quality and are underemployed. This tradeoff between flexibility and wages is a problem that cannot be explained away by a compensating differentials model. Because jobs offering flexibility are not distributed evenly across the occupational spectrum, the end result is the under-utilization of a material share of all of labor market participants.

Even though the new generation of women is relatively well-off in terms of match quality vis-à-vis its precursors, we do not in fact observe an overall improvement in occupational matches. Younger men share more of the burden, having worse matches than their childless counterparts and their older counterparts with children. Some of this shift-in-burden effect may be explained by the increase in the number of female breadwinners and broad change in gender roles. We have directly addressed the former but not the latter phenomenon in the present treatment.

Moreover, women in particular may be dropping out of the labor market or choosing part-time employment which may ultimately if not immediately cause them to drop out of our analysis sample. Appendix 3 shows that among individuals who have a child, those who were employed had remained in employment more than 90 percent of the time in the case of males and at least 80 percent of the time for females. However, some of the former and most of the latter did so in jobs that were not full time. If worse-matched individuals are precisely those who are dropping out, what we estimate as the gender differences in mismatch in the present paper is probably a lower bound to the true effect.

References

Altonji, Joseph. G., and Robert A. Shakotko. 1987. “Do Wages Rise with Job Seniority?” *Review of Economic Studies* 54(3): 437-459.

Autor, David H., and David Dorn. 2013. "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review* 103(5): 1553-1597.

Battu, Harminder, Clive R. Belfield, and Peter J. Sloane. 2000. “How Well Can We Measure Graduate Over-Education and Its Effects?” *National Institute Economic Review* No. 71: 82-93.

BLS. 2017. “Women in the Labor Force: A Databook.” Report 1065. Washington, DC: Bureau of Labor Statistics. Available at: https://www.bls.gov/opub/reports/womens-databook/2016/pdf/home.pdf.

Böheim, René, Dominik Grübl, and Mario Lackner. 2017. Differences in Competitiveness.” Ifo DICE Report 15(2): 13-17. Munich: ifo Institute-Leibniz Institute for Economic Research at the University of Munich.

Deming David. J. 2017a. “The Growing Importance of Social Skills in the Labor Market.” *Quarterly Journal of Economics* 132 (4): 1593-1640.

Deming, David, 2017b. “Replication Data for: “The Growing Importance of Social Skills in the Labor Market””, doi:10.7910/DVN/CYPKZH, Harvard Dataverse, V1.

Dorn, David. 2009. “Essays on Inequality, Spatial Interaction, and the Demand for Skills." Dissertation/University of St. Gallen, No. 3613, September.

Frank, Robert. 1978. “Why Women Earn Less: The Theory and Estimation of Differential Overqualification" *American Economic Review,* 68(June): 360-73

Goldin, Claudia. 2014. “A Grand Gender Convergence: Its Last Chapter.” *American Economic Review* 104(4): 1091-1119.

Guvenen, Fatih, Burhan Kuruscu, Satoshi Tanaka, and David Wiczer. 2016. “Multidimensional Skill Mismatch.” Mimeographed, University of Minnesota.

Gneezy, Uri, and John A. List. 2103. *The Why Axis: Hidden Motives and the Undiscovered Economics of Everyday Life.* New York: Public Affairs.

Hartog, Joop. 2000. “Over-Education and Earnings: Where are We, Where Should We Go?” *Economics of Education Review* 19(2): 149-159.

Hirsch B.T and Julia Manzella. 2015. “Who Cares – and Does it Matter? Measuring Wage Penalties for Caring Work”, in Solomon W. Polachek, Konstantinos Tatsiramos, Klaus F. Zimmermann (ed.) *Gender Convergence in the Labor Market (Research in Labor Economics, Volume 41)* Emerald Group Publishing Limited, pp.213 - 275

Johansson, Mats, and Katarina Katz. 2007. “Wage Differences between Women and Men in Sweden – The Impact of Skill Mismatch.” Working Paper 2007: 13, Institute for Labor Market Policy Evaluation/IFAU, Uppsala, Sweden.

Kahn, Lawrence B. 2010. “The Long-Term Consequences of Graduating from College in a Bad Economy.” *Labour Economics* 17(2): 303-316.

Lise, Jeremy, and Fabien Postel-Vinay. 2016. "Multidimensional Skills, Sorting, and Human Capital Accumulation." Mimeographed, University College London

Liu, Kai, Kjell G. Salvanes, and Erik Ø. Sørensen. 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.

Mavromaras, Kostas, Seamus McGuinness, Nigel O'Leary, Peter Sloane, and Zhang Wei. 2013. "Job Mismatches and Labour Market Outcomes: Panel Evidence on University Graduates." *Economic Record* 89 (286): 382–395.

Niederle, Muriel, and Lise Vesterlund. 2007. "Do Women Shy Away from Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics* 122 (30: 1067-1101.

Niederle, Muriel, and Lise Vesterlund. 2010. "Exploring the Gender Gap in Math Test Scores: The Role of Competition." *Journal of Economic Perspectives* 24(2): 129-144.

Robst, John. 2007. "Education, College Major, and Job Match: Gender Differences in Reasons for Mismatch." *Education Economics* 15(2): 159-175.

Summerfield, Fraser. 2015. "Matching Skill and Tasks: Cyclical Fluctuations in the Overqualification of New Hires." Mimeographed, University of Aberdeen.

Yamaguchi, Shintaro. 2010. "The Effect of Match Quality and Specific Experience on Career Decisions and Wage Growth." *Labour Economics* 17: 407-423.

Table 1. Sample Construction

	Criterion for sample selection	NLSY79 (1979-2014) Remaining Individuals Male	Female	Remaining Observations Male	Female	NLSY97 (1997-2013) Remaining Individuals Male	Female	Remaining Observations Male	Female
0	Entire Sample	6,403	6,283	166,478	163,358	4,599	4,385	73,584	70,160
1	In the Cross-sectional sample/ not oversampled	3,003	3,108	78,078	80,808	3,459	3,289	55,344	52,624
2	Not working before data sample period	2,394	2,703	62,244	70,278	3,276	3,172	52,416	50,752
3	Worked more than 1200 hours for the last 2 years	2,254	2,440	35,647	32,151	2,707	2,554	20,496	17,793
4	Not in the military for 2 years or more	2,253	2,440	33,828	30,161	2,707	2,554	20,492	17,793
5	Not in school	2,155	2,225	28,334	22,348	2,315	2,086	15,106	11,376
6	Currently Working	2,154	2,223	28,262	22,295	2,313	2,085	15,067	11,367
7	Have valid occupation and industry information	2,146	2,217	26,645	21,144	2,288	2,045	13,356	10,099
8	Older than 16	2,146	2,217	26,645	21,144	2,288	2,045	13,356	10,098
9	Have valid ASVAB scores and Ability Measurements	1,898	1,996	24,280	19,558	1,653	1,526	10,310	7,917
10	Have valid wage information	1,890	1,980	23,261	18,761	1,601	1,443	9,060	6,949
11	Have no missing information on variables of interest	1,890	1,980	23,261	18,761	1,588	1,434	8,981	6,912

Notes: We are using annual data and not the monthly job arrays.

Table 2. Descriptive Statistics of the Sample

		NLSY79			NLSY97		
Variable	**Definition**	**All**	**Male**	**Female**	**All**	**Male**	**Female**
Female	0/1 Dummy (=1 if female)	0.45			0.43		
Age at date of interview	Age in years	33.8	33.5	34.2	25.1	25.0	25.3
Less than high school	0/1 Dummy, less than 12 years of education	0.06	0.08	0.05	0.16	0.18	0.13
High school	0/1 Dummy, 12 years of education	0.45	0.46	0.45	0.33	0.37	0.28
Some college	0/1 Dummy, some years in collge	0.19	0.17	0.22	0.21	0.21	0.22
College or more	0/1 Dummy, at least 4-year college graduate	0.29	0.29	0.28	0.30	0.24	0.38
African-American	0/1 Dummy	0.11	0.11	0.12	0.14	0.13	0.14
Hispanic	0/1 Dummy	0.06	0.07	0.06	0.13	0.13	0.13
Have at least one child	0/1 Dummy	0.49	0.43	0.57	0.29	0.24	0.36
Have at least one child (age <=33)	0/1 Dummy (NLSY79, age <33)	0.34	0.30	0.39			
Age at the first birth	Age in years	26.5	27.5	25.3	24.0	24.2	23.7
Age at the first birth (age <=33)	Age in years (NLSY79, age < 33)	24.9	25.6	24.0			
Single	0/1 Dummy	0.28	0.31	0.25	0.67	0.70	0.63
Ever married	0/1 Dummy, (married, divorced, widowed, seperated)	0.72	0.69	0.75	0.33	0.30	0.37
Age at the first marriage	Age in years	24.8	25.7	23.7	24.7	25.0	24.4
Age at the first marriage (age <=33)	Age in years (NLSY79 age < 33)	23.5	24.3	22.6			
Traditional male occupation	< 20% female in both 1980 and 2000	0.23	0.38	0.04	0.23	0.39	0.03
Traditional female occupation	> 80% female in both 1980 and 2000	0.15	0.03	0.31	0.14	0.04	0.26
Traditional mixed occupation	20% to 80% female in both 1980 and 2000	0.51	0.48	0.53	0.51	0.46	0.57
Average labor market experience	Mean years worked	14.3	14.4	14.2	8.75	8.67	8.85
Median labor market experience	Median years worked	13.0	13.0	12.0	9.00	9.00	9.00
Average occupational tenure	Mean years worked in the same occupation	6.36	6.35	6.37	3.52	3.50	3.54
Median occupational tenure	Median years worked in the same occupation	4.00	4.00	4.00	3.00	3.00	3.00

Table 3. Descriptive Statistics of Workers' Skill Endowments

		NLSY79					NLSY97				
		Number of Individuals	Math	Verbal	STM	Social	Number of Individuals	Math	Verbal	STM	Social
Whole sample	Male	1890	0.55	0.52	0.61	0.56	1588	0.51	0.49	0.56	0.49
			[0.30]	[0.30]	[0.30]	[0.22]		[0.29]	[0.29]	[0.30]	[0.22]
	Female	1980	0.50	0.53	0.42	0.55	1434	0.52	0.52	0.46	0.56
			[0.28]	[0.29]	[0.25]	[0.23]		[0.28]	[0.28]	[0.26]	[0.22]
Less than high school	Male	149	0.20	0.19	0.30	0.43	268	0.28	0.28	0.35	0.45
			[0.16]	[0.18]	[0.26]	[0.22]		[0.22]	[0.23]	[0.27]	[0.22]
	Female	109	0.21	0.20	0.18	0.36	187	0.25	0.28	0.24	0.48
			[0.18]	[0.20]	[0.16]	[0.18]		[0.21]	[0.22]	[0.21]	[0.23]
High school	Male	780	0.41	0.38	0.51	0.52	455	0.39	0.39	0.48	0.48
			[0.25]	[0.26]	[0.29]	[0.21]		[0.25]	[0.26]	[0.28]	[0.21]
	Female	786	0.36	0.42	0.33	0.50	316	0.36	0.36	0.34	0.54
			[0.23]	[0.25]	[0.21]	[0.22]		[0.23]	[0.23]	[0.22]	[0.21]
Some college	Male	358	0.57	0.55	0.65	0.60	354	0.50	0.49	0.56	0.48
			[0.26]	[0.26]	[0.26]	[0.28]		[0.26]	[0.27]	[0.29]	[0.21]
	Female	459	0.49	0.54	0.41	0.55	309	0.50	0.53	0.46	0.56
			[0.24]	[0.26]	[0.22]	[0.22]		[0.24]	[0.25]	[0.24]	[0.21]
College or more	Male	603	0.81	0.75	0.81	0.63	511	0.73	0.70	0.74	0.54
			[0.19]	[0.20]	[0.19]	[0.21]		[0.22]	[0.23]	[0.23]	[0.22]
	Female	626	0.73	0.74	0.58	0.64	622	0.68	0.68	0.59	0.60
			[0.22]	[0.22]	[0.22]	[0.22]		[0.23]	[0.23]	[0.23]	[0.21]

Notes: Average percentile ranks are reported. Educational groups are defined by the highest degree ever completed. Standard errors are reported in brackets.

Table 4. The Type and Magnitude of Mismatch, by Experience, Skill Type, and Gender (NLSY79 and NSLY97)

			[1]	[2]	[3]	[4]	[5]	[6]
			Magnitude of Over-qualification	**Percent with Endowment > Requirement**	**Share of Over-qualified**	**Magnitude of Under-qualification**	**Percent with Endowment < Requirement**	**Share of Under-qualified**
					NLSY79			
Experience <= 10 years	All Skills	Male	0.200	-	50.7%	0.067	-	21.8%
		Female	0.193	-	50.7%	0.071	-	22.5%
	Math	Male	0.192	66.3%	39.0%	0.062	34.1%	14.6%
		Female	0.193	68.0%	40.1%	0.064	32.1%	14.0%
	Verbal	Male	0.197	33.2%	37.8%	0.056	66.8%	12.3%
		Female	0.197	33.4%	40.8%	0.065	66.6%	12.3%
	Social	Male	0.231	74.4%	48.6%	0.048	25.6%	13.1%
		Female	0.218	70.4%	45.1%	0.064	29.6%	16.3%
	STM	Male	0.180	40.9%	35.8%	0.103	59.1%	21.4%
		Female	0.164	38.1%	34.5%	0.089	61.9%	18.0%
Experience > 10 years	All Skills	Male	0.172	-	44.4%	0.082	-	26.5%
		Female	0.160	-	42.1%	0.093	-	29.5%
	Math	Male	0.170	64.8%	34.8%	0.073	35.2%	17.0%
		Female	0.166	63.2%	34.6%	0.085	36.8%	19.5%
	Verbal	Male	0.164	39.0%	31.1%	0.077	61.0%	17.4%
		Female	0.159	39.5%	32.8%	0.085	60.5%	20.5%
	Social	Male	0.170	61.0%	35.1%	0.084	39.0%	22.5%
		Female	0.166	59.0%	34.5%	0.102	41.0%	25.7%
	STM	Male	0.183	37.7%	36.3%	0.092	62.3%	18.5%
		Female	0.149	42.6%	31.5%	0.102	57.4%	21.0%
					NLSY97			
Experience <= 10 years	All Skills	Male	0.181	-	48.8%	0.084	-	32.1%
		Female	0.180	-	48.2%	0.078	-	27.2%
	Math	Male	0.167	60.5%	33.6%	0.078	39.5%	23.3%
		Female	0.176	63.9%	36.6%	0.070	36.1%	20.5%
	Verbal	Male	0.195	35.0%	38.6%	0.063	65.0%	17.8%
		Female	0.163	41.8%	31.8%	0.087	58.2%	24.4%
	Social	Male	0.198	69.6%	42.4%	0.062	30.4%	15.2%
		Female	0.194	65.1%	40.7%	0.077	34.9%	20.2%
	STM	Male	0.164	45.8%	30.5%	0.133	54.2%	28.5%
		Female	0.189	36.9%	37.6%	0.076	63.1%	17.2%
Experience > 10 years	All Skills	Male	0.181	-	47.7%	0.088	-	33.9%
		Female	0.169	-	45.5%	0.091	-	33.4%
	Math	Male	0.178	63.4%	37.5%	0.073	36.6%	20.6%
		Female	0.184	67.0%	38.9%	0.068	33.0%	19.1%
	Verbal	Male	0.187	37.2%	35.1%	0.073	62.8%	20.1%
		Female	0.143	45.1%	28.0%	0.097	54.9%	27.0%
	Social	Male	0.162	59.7%	34.5%	0.100	40.3%	24.7%
		Female	0.148	53.4%	31.8%	0.123	46.6%	31.0%
	STM	Male	0.195	38.4%	36.0%	0.104	61.6%	23.2%
		Female	0.199	35.9%	40.8%	0.077	64.1%	17.7%

Notes: "Magnitude of Over-qualification" is an average non-standardized measure of the distance between the worker's endowments and occupational requirements, when the worker's average skill endowment exceeds the average occupational skill requirements. "Endowment > Requirement" is a crude definition of over-qualification, here the percentage of workers with endowments that are greater than the skill levels required by their occupation. "Share of over-qualified" on the other hand gives the share of workers who are more than one standard deviation more endowed than required by the occupation. Measures of under-qualification are similarly constructed.

Table 5. The Determinants of Mismatch (NLSY79)

	(1)	(2)	(3)	(4)
Female	-0.0253	-0.0309	-0.0325	-0.031
	[0.0256]	[0.0278]	[0.0277]	[0.0279]
Female* College or more		0.1536	0.1509	0.1557
		[0.0509]**	[0.0508]**	[0.0510]**
College or more		-0.2624	-0.2649	-0.2642
		[0.0388]**	[0.0389]**	[0.0388]**
Observations	42022	42022	42022	42022

Notes: The dependent variable is the standardized total amount of mismatch. All specifications are estimated using OLS-based models. The first column includes only the female dummy. Column (2) also controls for demographic/human capital characteristics (race, completed years of schooling), average measures of individual's skills and occupational requirements, and year dummies. Column (3) adds further controls for employer tenure, occupational tenure, and experience, as well as their polynomials; also included are the interaction term of skills and occupational tenure and the interaction term of occupational requirements and occupational tenure. The set of variables in column (3) will henceforth will be referred to as the full set of controls. In column (4) employer tenure, occupational tenure, and total experience, as well as their quadratic forms and interaction terms, are instrumented. Clustered standard errors are reported in brackets. **, *, + indicate significance at the 0.01, 0.05 and 0.1 levels, respectively.

Table 6. The Determinants of Mismatch: The Role of Fertility and Marriage (NLSY79)

	IV			FE			IV-FE		
	All	**Male**	**Female**	**All**	**Male**	**Female**	**All**	**Male**	**Female**
Have at least one child	-0.0307	-0.0129	0.076	-0.0189	-0.0156	0.1138	-0.0342	-0.0241	0.0965
	[0.0292]	[0.0294]	[0.0350]*	[0.0159]	[0.0157]	[0.0225]**	[0.0161]*	[0.0159]	[0.0227]**
Female* Have at least one child	0.0971			0.1305			0.1393		
	[0.0433]*			[0.0249]**			[0.0250]**		
Female	-0.0543								
	[0.0383]								
Ever married	-0.0273	-0.0047	-0.0125	0.0138	0.0162	-0.0455	0.0056	0.012	-0.0518
	[0.0310]	[0.0316]	[0.0381]	[0.0183]	[0.0185]	[0.0238]+	[0.0184]	[0.0186]	[0.0240]*
Female * Ever married	0.034			-0.0576			-0.0492		
	[0.0477]			[0.0277]*			[0.0279]+		
Observations	42022	23261	18761	42022	23261	18761	42022	23261	18761

Notes: The dependent variable is the standardized total amount of mismatch. In the IV specifications, employer tenure, occupational tenure and total experience, as well as their quadratic forms and interaction terms are instrumented. All specifications use the full set of controls is listed before in Table 5. Robust standard errors are reported in brackets. **, *, + indicate significance at the 0.01, 0.05 and 0.1 levels, respectively.

Table 7. The Probability of Being Over-qualified: The Role of Fertility and Marriage (NLSY79)

	Male		Female	
	Coefficient	**Odds Ratio**	**Coefficient**	**Odds Ratio**
Have at least one child	-0.0559	0.95	0.6675	1.95
	[0.2403]		[0.3574]+	
Ever married	-0.0403	0.96	0.5133	1.67
	[0.2855]		[0.3734]	
Observations	15149		11125	

Notes: The dependent variable is a dummy that equals 1 when the individual's average total amount of skill surplus exceeds 1 standard deviation. Logit FE estimates are reported. The sample size on this table is smaller because observations with all positive or all negative outcomes are dropped. The baseline group are those childless individuals who never get married. All specifications have the full set of controls. The standard errors for the pooled logit estimation are clustered at individual level. Standard errors are reported in brackets. **, *, + indicate significance at the 0.01, 0.05 and 0.1 levels, respectively.

Table 8. Mismatch and the Fertility Timeline (NLSY79)

	OLS		FE		IV-FE	
	Male	**Female**	**Male**	**Female**	**Male**	**Female**
More than 3 Years before the first birth	0.0112	0.0474	0.0361	0.0261	0.0418	0.0349
	[0.0403]	[0.0444]	[0.0237]	[0.0292]	[0.0238]+	[0.0295]
0-6 Years after the first birth	-0.0429	0.0567	-0.0375	0.0962	-0.0442	0.0829
	[0.0324]	[0.0393]	[0.0202]+	[0.0263]**	[0.0204]*	[0.0265]**
More than 6 Years after the first birth	-0.003	0.1043	0.015	0.1731	0.0022	0.1512
	[0.0391]	[0.0405]*	[0.0243]	[0.0298]**	[0.0246]	[0.0302]**
Observations	23261	18761	23261	18761	23261	18761

Notes: The dependent variable is the standardized total amount of mismatch. The baseline group is individuals 0-3 years before the first birth. In the IV-FE specification, employer tenure, occupational tenure and total experience, as well as their quadratic forms and interaction terms are instrumented. All models include the full set of controls. Robust standard errors are reported in brackets. **, *, + indicate significance at the the 0.01, 0.05 and 0.1 levels, respectively.

Table 9. Mismatch and Fertility Delay (NLSY79)

	OLS		IV-1		IV-2	
Measure of age at first birth	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**
Relative age at the first birth	0.0038	-0.004	0.0088	-0.0398	0.0075	-0.0361
	[0.0013]**	[0.0015]**	[0.0090]	[0.0166]*	[0.0092]	[0.0160]*
Absolute age at the first birth	0.0031	-0.0038	0.0087	-0.0366	0.0075	-0.0332
	[0.0013]*	[0.0017]*	[0.0090]	[0.0151]*	[0.0091]	[0.0146]*

Notes: The dependent variable is the standardized amount of mismatch. Age at birth is measured in two ways. The relative birth age is calculated as the difference between the *the year of the first birth* and *the year one enters the labor market.* All specifications have the full set of controls. In IV-1, individuals' relative/absolute birth ages are instrumented using their siblings' average age at the first birth. In IV-2, individuals' relative/absolute birth ages, employer tenure, occupational tenure and labor market experience are all instrumented. Robust standard errors are reported in brackets.**, *, + indicate significance at the 0.01, 0.05 and 0.1 levels, respectively.

Table 10. Mismatch, Fertility, and Occupational Flexibility (NLSY79)

	Male	Female	Male	Female
Have at least one child	-0.0199	0.1185		
	[0.0160]	[0.0235]**		
Have at least one child * Flexibility score	0.0409	0.0636		
	[0.0137]**	[0.0173]**		
Flexibility score	-0.0247	-0.0238	-0.0365	-0.0236
	[0.0094]**	[0.0132]+	[0.0126]**	[0.0163]
More than 3 years before the first birth			0.0444	0.03
			[0.0239]+	[0.0313]
0-6 years after the first birth			-0.0411	0.0918
			[0.0205]*	[0.0280]**
More than 6 years after the first birth			0.0056	0.178
			[0.0246]	[0.0311]**
More than 3 years before the first birth * Flexibility score			0.0355	-0.0104
			[0.0220]	[0.0275]
0-6 years after the first birth * Flexibility score			0.0415	0.0282
			[0.0183]*	[0.0254]
More than 6 years after the first birth * Flexibility score			0.0456	0.0766
			[0.0172]**	[0.0211]**
Observations	23261	18761	23261	18761

Notes: In the above specifications, the dependent variable is the standardized total amount of mismatch. All models include the full set of controls. The FE model estimates are not reported because of a space constraint. They are almost identical to the IV-FE results. Standard errors are reported in brackets. **, *, + indicate significance at the 0.01, 0.05 and 0.1 levels, respectively.

Table 11. Mismatch and Wage Outcomes (NLSY79)

	All			Male			Female			All	Male	Female
Mismatch	-0.0236	-0.0213	-0.0038	-0.0243	-0.0208	-0.0035	-0.0227	-0.0219	-0.0083	0.0049	0.0062	0.0038
	[0.0031]**	[0.0053]**	[0.0056]	[0.0033]**	[0.0071]**	[0.0072]	[0.0030]**	[0.0067]**	[0.0067]	[0.0060]	[0.0074]	[0.0071]
Mismatch*Occupation tenure		-0.0006	-0.0005		-0.001	-0.0008		-0.0002	-0.0005	-0.0009	-0.0006	-0.0012
		[0.0014]	[0.0014]		[0.0019]	[0.0019]		[0.0017]	[0.0017]	[0.0014]	[0.0020]	[0.0018]
Cumulative Mismatch			-0.0598			-0.0606			-0.0526	-0.0421	-0.0384	-0.0509
			[0.0036]**			[0.0040]**			[0.0038]**	[0.0043]**	[0.0047]**	[0.0042]**
Female	-0.1429	-0.1416	-0.1449							-0.1257		
	[0.0096]**	[0.0096]**	[0.0118]**							[0.0119]**		
Female*Mismatch	0.0014	0.0056	0.0019							0.0063		
	[0.0042]	[0.0052]	[0.0054]							[0.0064]		
Female*Mismatch*Occupation tenure		-0.0014	-0.0015							-0.0013		
		[0.0010]	[0.0010]							[0.0012]		
Female*Cumulative mismatch			0.005							-0.0052		
			[0.0043]							[0.0051]		
Mismatch*College and above										-0.0311	-0.0362	-0.0429
										[0.0090]**	[0.0095]**	[0.0086]**
Mismatch*Occupation Tenure*College and above										0.0014	-0.0004	0.0029
										[0.0017]	[0.0018]	[0.0015]+
Cumulative Mismatch*College and above										-0.0513	-0.0639	-0.0217
										[0.0065]**	[0.0075]**	[0.0073]**
Female*Mismatch*College and above										-0.0172		
										[0.0117]		
Female*Mismatch*Occupation Tenure*College and above										-0.0008		
										[0.0022]		
Female*Cumulative Mismatch*College and above										0.0115		
										[0.0073]		
College and Above										0.4169	0.4812	0.3331
										[0.0167]**	[0.0237]**	[0.0232]**
Observations	42022	42022	42022	23261	23261	23261	18761	18761	18761	42022	23261	18761

Notes: IV model estimates are reported. All specifications use the full set of controls. Occupational tenure variables and continuing job dummy are instrumented as described in the text. Robust standard errors are reported in brackets. **, *, + indicate significance at the 0.01, 0.05, and 0.1 levels, respectively.

Table 12. Mismatch and Gender Wage Gap, by Experience, Early Match Quality and Education

		Male		Male: Wage Effect		Female		Female: Wage Effect		Gender Gap				
		Cumulative Mismatch	Current Mismatch	Cumulative Mismatch	Current Mismatch	Cumulative Mismatch	Current Mismatch	Cumulative Mismatch	Current Mismatch	Cumulative Mismatch	Current Mismatch	Occupational Tenure (insignificant)	Total (significant coefficients only)	Total (signficiant and insignificant coefficients)
College Graduates														
Experience and Early Match Quality	10 Years Exp.													
	Best	1.53	1.29	-0.14	-0.04	1.87	2.14	-0.17	-0.07	-0.03	-0.03	0.003	-0.058	-0.055
	Worst	3.71	2.05	-0.35	-0.06	3.65	2.05	-0.34	-0.06	0.01	0.00	0.000	0.006	0.006
	20 Years Exp.													
	Best	1.8	1.31	-0.17	-0.04	2.36	2.01	-0.22	-0.06	-0.05	-0.02	0.006	-0.074	-0.068
	Worst	3.41	1.94	-0.32	-0.06	3.57	1.83	-0.33	-0.06	-0.01	0.00	-0.001	-0.012	-0.012
	30 Years Exp.													
	Best	1.92	1.12	-0.18	-0.03	2.56	1.56	-0.24	-0.05	-0.06	-0.01	0.005	-0.073	-0.068
	Worst	3.37	1.84	-0.31	-0.06	3.57	2.11	-0.33	-0.07	-0.02	-0.01	0.003	-0.027	-0.024
Non-College Graduates														
Experience and Early Match Quality	10 Years Exp.													
	Best	1.56	1.43	-0.07	0.01	1.6	1.59	-0.07	0.01	-0.002	0.00	-0.002	-0.002	-0.002
	Worst	3.96	2.53	-0.17	0.01	3.98	2.91	-0.17	0.01	-0.001	0.00	-0.004	-0.001	-0.003
	20 Years Exp.													
	Best	2.01	1.68	-0.08	0.01	2.07	1.76	-0.09	0.01	-0.003	0.00	-0.002	-0.003	-0.004
	Worst	4.08	2.38	-0.17	0.01	4.13	2.54	-0.17	0.01	-0.002	0.00	-0.003	-0.002	-0.005
	30 Years Exp.													
	Best	2.17	1.43	-0.09	0.01	2.23	1.56	-0.09	0.01	-0.003	0.00	-0.004	-0.003	-0.006
	Worst	4.11	2.32	-0.17	0.01	4.17	2.52	-0.18	0.01	-0.003	0.00	-0.006	-0.003	-0.008

Notes: Early match quality is the rank according to an individual's match quality over the first 5 years of experience. In this table we have the bottom decile (Best) and the top decile (Worst) of the mismatch distribution. Current and past mismatch values are averages for these groups with the given level of experience and early career match quality. The occupational tenure-mismatch interaction effect is not significant in Table 11. The last column in this table present the results using all mismatch related coefficients, both significant and insignificant, while the penultimate column considers only those coefficients that are precisely estimated.

Table 13. Determinants of Mismatch and Over-qualification (Cohort Differences)

	Mismatch		Over-qualified	
	OLS	IVREG	PROBIT	IVPROBIT
Female	-0.033	-0.0386	-0.1607	-0.1569
	[0.0141]*	[0.0142]**	[0.0385]**	[0.0405]**
Female* NLSY97	-0.0776	-0.0738	-0.128	-0.1238
	[0.0225]**	[0.0225]**	[0.0633]*	[0.0636]+
Female* College or more	0.0984	0.102	0.0896	0.066
	[0.0280]**	[0.0281]**	[0.0736]	[0.0801]
Female*College or more* NLSY97	0.0333	0.0328	0.111	0.125
	[0.0427]	[0.0429]	[0.1133]	[0.1208]
NLSY97*College or more	0.2059	0.181	0.2359	0.2513
	[0.0298]**	[0.0304]**	[0.0816]**	[0.0845]**
College or more	-0.2489	-0.2304	-0.2031	-0.2145
	[0.0202]**	[0.0204]**	[0.0571]**	[0.0583]**
NLSY97	0.0952	0.1029	0.2735	0.1971
	[0.0146]**	[0.0181]**	[0.0454]**	[0.0513]**
Observations	39112	39112	39112	39112

Notes: See the notes to Table 5 and 7 for definition of outcome variables. NLSY97 is a dummy which equals 1 if the observation is from the NSLY97 cohort.

Table 14. Determinants of Mismatch: The Roles of Family, Fertility and Cohort

	All	Male	Female	All	Male	Female
Have at least one child	-0.0322	-0.0261	0.0842			
	[0.0193]+	[0.0198]	[0.0284]**			
Have at least one child*NLSY97	0.0441	0.0564	-0.0803			
	[0.0322]	[0.0330]+	[0.0438]+			
Female*Have at least one child	0.1209					
	[0.0309]**					
Female*Have at least one child*NSLY97	-0.1107					
	[0.0506]*					
More than 3 Years before the first birth				0.038	0.0379	0.0353
				[0.0200]+	[0.0263]	[0.0307]
0-6 Years after the first birth				0.007	-0.0195	0.0588
				[0.0181]	[0.0230]	[0.0295]*
More than 6 Years after the first birth				0.1038	0.0667	0.1635
				[0.0279]**	[0.0358]+	[0.0442]**
More than 3 Years before the first birth*NLSY97				0.0151	0.0007	0.0472
				[0.0380]	[0.0417]	[0.0482]
0-6 Years after the first birth*NLSY97				0.0106	0.0519	-0.0276
				[0.0353]	[0.0379]	[0.0463]
More than 6 Years after the first birth*NLSY97				-0.0911	-0.0339	-0.0825
				[0.0552]+	[0.0591]	[0.0685]
Female*More than 3 Years before the first birth*NLSY97				0.0134		
				[0.0476]		
Female*0-6 Years after the first birth*NLSY97				0.0224		
				[0.0461]		
Female*More than 6 Years after the first birth*NLSY97				0.1041		
				[0.0706]		
Observations	39112	22152	16960	39112	22152	16960

Notes: Estimates from IV-FE models are reported. See also the notes to Tables 5, 6 and 13.

Table 15. Skill Mismatch, Fertility, and Breadwinners

NLSY79

	IV		FE		FE-IV	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Breadwinner*Have at least one child	0.0951	-0.0306	0.0512	-0.0738	0.0546	-0.0735
	[0.0502]+	[0.0649]	[0.0331]	[0.0390]+	[0.0333]	[0.0392]+
Breadwinner	-0.08	-0.1016	-0.0763	0.0145	-0.0777	0.0107
	[0.0415]+	[0.0478]*	[0.0281]**	[0.0306]	[0.0282]**	[0.0307]
Have at least one child	0.0276	0.0226	0.0824	0.0517	0.0569	0.0468
	[0.0468]	[0.0646]	[0.0315]**	[0.0391]	[0.0318]+	[0.0393]
Observations	14039	15998	14039	15998	14039	15998

NLSY97

Breadwinner*Have at least one child	-0.0301	0.157	0.0014	0.0184	0.0045	0.0182
	[0.0926]	[0.1230]	[0.0614]	[0.0754]	[0.0619]	[0.0757]
Breadwinner	-0.0153	-0.0687	-0.0878	-0.07	-0.0835	-0.0666
	[0.0661]	[0.0829]	[0.0453]+	[0.0566]	[0.0457]+	[0.0568]
Have at least one child	0.0431	-0.2095	0.0319	-0.0002	0.0365	0.0081
	[0.0734]	[0.1182]+	[0.0564]	[0.0736]	[0.0570]	[0.0740]
Observations	2282	2384	2282	2384	2282	2384

Notes: All specifications include full set of controls. The "breadwinners" are determined when an individual's spouse does not have any wage income or spousal wage income is less than that of the respondent.

Table16. Determinants Mismatch and Over-qualification, by Skill Type (NLSY79)

A. Mismatch

	Math Skills			**Verbal Skills**			**STM Skills**			**Social Skills**		
	All	**Male**	**Female**	**All**	**Male**	**Female**	**All**	**Male**	**Female**	**All**	**Male**	**Female**
Have at least one child	0.0429	0.0418	0.0949	-0.0057	-0.0155	0.084	0.019	0.0222	0.1107	0.0074	0.011	0.0827
	[0.0183]*	[0.0190]*	[0.0258]**	[0.0280]	[0.0181]	[0.0245]**	[0.0296]	[0.0192]	[0.0243]**	[0.0261]	[0.0168]	[0.0249]**
Have at least one child* College	-0.2347	-0.2228	-0.1379	-0.12	-0.1075	-0.1142	-0.1102	-0.1065	-0.0468	0.003	0.0126	-0.0652
or more	[0.0315]**	[0.0320]**	[0.0407]**	[0.0455]**	[0.0305]**	[0.0387]**	[0.0506]*	[0.0322]**	[0.0383]	[0.0513]	[0.0283]	[0.0394]+
Female* Have at least one child	0.0498			0.068			0.0894			0.0745		
	[0.0279]+			[0.0449]			[0.0449]*			[0.0471]		
Female* Have at least one	0.1063			0.0316			0.0621			-0.0494		
child*College or more	[0.0492]*			[0.0756]			[0.0768]			[0.0837]		
College or more	-0.131	-0.2215	-0.0159	-0.0792	-0.2367	0.1367	0.0321	0.0062	0.0683	0.0159	-0.2053	0.3146
	[0.0689]+	[0.0912]*	[0.1044]	[0.1113]	[0.0869]**	[0.0993]	[0.0925]	[0.0918]	[0.0982]	[0.1062]	[0.0806]*	[0.1010]**

B. Over-qualified

	Math Skills			**Verbal Skills**			**STM Skills**			**Social Skills**		
	All	**Male**	**Female**	**All**	**Male**	**Female**	**All**	**Male**	**Female**	**All**	**Male**	**Female**
Have at least one child	0.1992	0.2473	-0.2367	0.2912	0.2966	0.0339	0.165	0.1223	0.2694	-0.1728	-0.0979	0.329
	[0.1673]	[0.1946]	[0.2108]	[0.1774]	[0.1908]	[0.2175]	[0.1177]	[0.1185]	[0.1964]	[0.1314]	[0.1352]	[0.1908]+
Have at least one child* College	-0.5384	-0.8614	0.5685	-0.3057	-0.3034	0.1596	-0.0842	-0.1537	0.0999	-0.2732	-0.0886	-1.0051
or more	[0.2612]*	[0.2945]**	[0.3268]+	[0.2881]	[0.2981]	[0.3530]	[0.1807]	[0.1766]	[0.3023]	[0.2239]	[0.2223]	[0.3444]**
Female* Have at least one child	-0.3816			-0.2719			0.0112			0.6729		
	[0.2551]			[0.2492]			[0.1902]			[0.1988]**		
Female* Have at least one	0.9308			0.4983			0.0592			-0.3224		
child*College or more	[0.4244]*			[0.4383]			[0.3183]			[0.3821]		
College or more	1.0957	0.7753	0.8185	-0.953	-1.0511	-0.7854	0.0462	0.2553	-0.4793	0.5769	0.5895	0.6634
	[0.6998]	[1.1959]	[0.8384]	[0.7020]	[0.9722]	[1.0054]	[0.4327]	[0.5031]	[0.8801]	[0.6086]	[0.7705]	[0.9740]

Notes: In the upper panel the dependent variable is the standardized total amount of mismatch for any skill. The estimations are based on IV-Fixed Effect regressions in which employer tenure, occupational tenure, and total experience, as well as their quadratic forms and interaction terms, are instrumented. For the lower panel, a dummy indicator of over-qualification is the dependent variable. FE-Logit estimates are reported for this panel. All specifications use the full set of controls is listed before in Table 3. Robust standard errors are reported in brackets. **, *, + indicate significance at the 1%, 5% and 10% levels, respectively.

Table 17. Mismatch and Wages, by Skill Type (NLSY79)

	IV									
	[1]		[2]		[3]		OLS		IV-FE	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
Math mismatch	-0.0141	-0.0073	-0.0122	0.0009	-0.0153	-0.0076	-0.0062	0.0022	-0.009	-0.0091
	[0.0044]**	[0.0039]+	[0.0072]+	[0.0066]	[0.0080]+	[0.0068]	[0.0080]	[0.0086]	[0.0079]	[0.0060]
Verbal mismatch	-0.0078	-0.0048	0.0027	0.0031	0.0152	0.0136	-0.0067	-0.023	-0.0015	0.0061
	[0.0041]+	[0.0039]	[0.0069]	[0.0075]	[0.0083]+	[0.0079]+	[0.0074]	[0.0085]**	[0.0081]	[0.0069]
STM mismatch	-0.0028	-0.0055	-0.0098	-0.0243	-0.0043	-0.0207	-0.0208	-0.0153	-0.0087	-0.0063
	[0.0035]	[0.0037]	[0.0058]+	[0.0067]**	[0.0059]	[0.0068]**	[0.0061]**	[0.0081]+	[0.0063]	[0.0065]
Social-noncognitive mismatch	0.0054	0.0037	0.0013	0.0006	-0.0027	-0.0049	0.0009	0.0145	-0.0053	-0.0009
	[0.0036]	[0.0030]	[0.0059]	[0.0057]	[0.0067]	[0.0060]	[0.0071]	[0.0065]*	[0.0072]	[0.0059]
Math mismatch * Occupation tenure			-0.0007	-0.0025	-0.0006	-0.0024	-0.0048	-0.0069	-0.001	-0.0015
			[0.0020]	[0.0016]	[0.0020]	[0.0016]	[0.0020]*	[0.0023]**	[0.0019]	[0.0011]
Verbal mismatch * Occupation tenure			-0.0036	-0.0026	-0.0034	-0.0026	-0.0008	0.0041	-0.0037	-0.0032
			[0.0021]+	[0.0021]	[0.0021]	[0.0021]	[0.0018]	[0.0022]+	[0.0020]+	[0.0014]*
STM mismatch * Occupation tenure			0.0023	0.0057	0.0024	0.0058	0.0031	0.0025	0.0018	0.0018
			[0.0015]	[0.0018]**	[0.0015]	[0.0018]**	[0.0017]+	[0.0021]	[0.0013]	[0.0014]
Social-noncognitive mismatch * Occupation tenure			0.0012	0.0007	0.0011	0.0008	-0.0014	-0.0043	0.0003	-0.0001
			[0.0016]	[0.0014]	[0.0016]	[0.0014]	[0.0019]	[0.0018]*	[0.0016]	[0.0011]
Cummulative math mismatch					0.0041	-0.0076	0.0181	0.0455	-0.0464	-0.0224
					[0.0067]	[0.0068]	[0.0071]*	[0.0082]**	[0.0139]**	[0.0132]+
Cummulative verbal mismatch					-0.0241	0.0136	-0.021	-0.0514	-0.0438	-0.0228
					[0.0071]**	[0.0079]+	[0.0069]**	[0.0080]**	[0.0152]**	[0.0129]+
Cummulative STM mismatch					-0.0135	-0.0207	-0.0019	0.0237	0.0081	0.000
					[0.0038]**	[0.0068]**	[0.0051]	[0.0072]**	[0.0094]	[0.0113]
Cummulative social-noncognitive mismatch					0.0085	-0.0049	-0.0098	-0.0032	-0.0095	-0.0215
					[0.0057]	[0.0060]	[0.0057]+	[0.0050]	[0.0104]	[0.0102]*
Observations	23261	18761	23261	18761	23261	18761	23261	18761	23261	18761

Notes: See the notes to Table 11 and 16.

Appendix 1. The Mapping of Occupational Codes

The 2002 Census Occupation Codes (COC) are first converted to 2000 COC and then mapped to the 3-digit occupation codes (occ1990dd) constructed in Dorn (2009). Specifically, respondents' 2000 COCs were mapped to occ1990dd using the crosswalks downloaded from http://www.ddorn.net/data.htm on Sep.24, 2015. It emerged that there were 11 occupations which were not worked by NLSY97 respondents, 21 occupations that could not be mapped to occ1990dd, and 2 occupations that were miscoded. After Dorn, we assigned the approximate 1990dd code to the 21 un-mapped 2000 COC occupations to minimize observation loss. The list of occupations that were manually mapped is as follows:

2000 COC	Occupation Name	Occ1990dd	Occupation Name
123	Statisticians	68	Mathematicians and statisticians
134	Biomedical engineers	59	Engineers and other professionals, n.e.c.
383	Fish and game wardens	427	Protective service, n.e.c.
416	Food preparation and serving related workers, all other	444	Miscellaneous food preparation and service workers
631	Pile-driver operators	599	Misc. construction and related occupations
521	Correspondence clerks	326	Correspondence and order clerks
650	Reinforcing Iron and Rebar Workers	597	Structural metal workers
705	Electrical and Electronics Installers and Repairers, Transportation Equipment	533	Repairers of electrical equipment, n.e.c.
802	Milling and Planing Machine Setters, Operators, and Tenders, Metal and Plastic	703	Lathe, milling, and turning machine operatives
812	Multiple Machine Tool Setters, Operators, and Tenders, Metal and Plastic	684	Other precision and craft workers
884	Semiconductor Processors	779	Machine operators, n.e.c.
911	Ambulance Drivers and Attendants, Except Emergency Medical Technicians	809	Taxi cab drivers and chauffeurs
950	Conveyor Operators and Tenders	889	Laborers, freight, stock, and material handlers, n.e.c
150	Mining and Geological Engineers, Including Mining Safety Engineers	59	Petroleum, mining, and geological engineers
194	Nuclear Technicians	235	Other science technicians
602	Animal breeders	479	Animal Breeders; "Animal caretakers, except farm
692	Roustabouts, oil and gas	616	Miners
693	Helpers--extraction workers	617	Other mining occupations
752	Commercial drivers	809	Taxi cab drivers and chauffeurs
973	Shuttle car operators	808	Bus drivers
974	Tank car, truck, and ship loaders	859	Stevedores and misc. material moving occupations
467	Not in 2000 COC	No Code	N/A
617	Not in 2000 COC	No Code	N/A

After mapping, the occupations are divided into 6 aggregate groups, using do-files downloaded from http://www.ddorn.net/data.htm on September 24, 2015. The six groups, which are also used by Autor and Dorn (2013) are: managerial and professional specialty; technical, sales, and administrative support; services; farming, forestry, and fishing; precision production, craft, and repair; and operators, fabricators, and laborers. We used 14 industry sector groups capturing public employees with a public administration/public sector dummy. The remaining 13 (private) industry/sector groups are agriculture, forestry, and fisheries; mining; construction; manufacturing (non-durable goods); manufacturing (durable goods); transportation, communications, and other public utilities; wholesale trade; retail trade; finance, insurance, and real estate; business and repair services; personal services; entertainment and recreation services; and professional and related services.

Appendix 2. The Mapping of ASVAB and O*NET Components

ASVAB COMPONENT	O*NET Knowledge/Skill/Ability	O*NET COMPONENT
Verbal		
Word Knowledge	Ability	Inductive Reasoning
Paragraph Comprehension	Ability	Written Comprehension
	Ability	Oral Comprehension
	Knowledge	English Language
	Skill	Reading Comprehension
Math		
Arithmetic Reasoning	Ability	Deductive Reasoning
Math Knowledge	Ability	Inductive Reasoning
	Ability	Written Comprehension
	Ability	Number Facility
	Ability	Mathematical Reasoning
	Ability	Information Ordering
	Knowledge	Mathematics
	Skill	Science
	Skill	Mathematics
Science and Mechanical		
General Science	Ability	Deductive Reasoning
Mechanical	Ability	Inductive Reasoning
Electronics Information	Ability	Written Comprehension
	Knowledge	Mechanical
	Knowledge	Biology
	Knowledge	Computers and Electronics
	Knowledge	Engineering and Technology
	Knowledge	Chemistry
	Knowledge	Physics
	Knowledge	Building and Construction
	Skill	Technology Design
	Skill	Science
	Skill	Installation
	Skill	Troubleshooting
	Skill	Equipment Selection
	Skill	Operation and Control

Appendix 3. Fertility and Labor Market Attachment

Employment to Employment Transitions

	Employed 0-3 Years before the First Birth	Share that Remained Employed after the First Birth		
		0-3 years	3-6 years	6-10 years
Female	1560	0.94	0.80	0.83
Male	1468	0.97	0.92	0.91
		Share that Remained Employed after the Last Birth		
		0-3 years	3-6 years	6-10 years
Female	1560	0.84	0.79	0.82
Male	1468	0.95	0.88	0.85

Full-time to Full-time Transitions

	Full-Time Employed 0-3 Years before the First Birth	Share that Remained Full-Time Employed after the First Birth		
		0-3 years	3-6 years	6-10 years
Female	1505	0.46	0.34	0.47
Male	1466	0.66	0.64	0.76
		Share that Remained Full-Time Employed after the Last Birth		
		0-3 years	3-6 years	6-10 years
Female	1505	0.43	0.42	0.54
Male	1466	0.74	0.70	0.73

Note: A person is defined as employed if that individual has a valid occ1990dd occupation and as full-time employed if, in addition, he or she works for more than 35 hours/week.