CESifo WORKING PAPERS
7350
2018
November 2018

Occupational Match Quality and Gender over Two Cohorts

John T. Addison, Liwen Chen, Orgul D. Ozturk

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editors: Clemens Fuest, Oliver Falck, Jasmin Gröschl
www.cesifo-group.org/wp

An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: www.CESifo-group.org/wp

CESifo *Working Paper No. 7350*
Category 4: Labour Markets

Occupational Match Quality and Gender over Two Cohorts

Abstract

Using a multi-dimensional measure of occupational mismatch, we report distinct gender differences in match quality and changes in match quality over the course of careers. A substantial portion of the gender wage gap stems from match quality differences among more educated individuals. College-educated females are significantly more mismatched than males. Individuals with children and in more flexible occupations also tend to be more mismatched. Again, this is especially true of women. Cohort effects are also discernible: college-educated males of the younger cohort have lower match quality than the older cohort, even as the new generation of women is doing better.

JEL-Codes: J130, J160, J220, J230, J240, J310, J330, J380.

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

John T. Addison
Darla Moore School of Business
University of South Carolina
1014 Greene Street
USA - Columbia, SC 29208
ecceaddi@moore.sc.edu

Liwen Chen
Faculty of Economics and Management
East China Normal University
3663 North Zhongshan Road
Shanghai, China 200062
liwenchen.china@gmail.com

Orgul D. Ozturk
Darla Moore School of Business
University of South Carolina
1014 Greene Street
USA - Columbia, SC 29208
odozturk@moore.sc.edu

1 Introduction

There is an influential literature that emphasizes the importance of quality of job matches for wage growth (Simpson 1990; Light and Ureta 1992; Loprest 1992). Finding the right skill match may require trying out different occupations, a process that implies costly search (Burdett 1978; Jovanovic 1979). This process may be more onerous and costly for workers who are constrained in their search options and flexibility to move. Traditionally, women have been expected to do more of the housework and care for dependents. They have moreover historically assumed a secondary role in family labor supply allocation decisions. Unlike their male counterparts, therefore, women may secure worse matches, leading us to anticipate greater mismatch in the labor market for women. In short, mismatch may be expected to explain some of the gender wage gap.

Most recent additions to this literature seek formally to model and estimate the quality of occupational matches (e.g. Lise and Postel-Vinay 2016) using multidimensional objective skill measures and occupational requirements along those dimensions. This approach would seem well suited to the study of wage disparities resulting from occupational gender segregation in the labor market. However, in studies linking wage deficits to imbalances in occupational match quality, women are distinguished by their absence. The present paper seeks to remedy this omission. It documents gender differences in occupational match quality, analyzes the role of life events such as marriage and fertility in this process, and quantifies the portion of the wage gap that may be attributable to disparities in match quality.

Although our main contribution is the analysis of gender differences in match quality over a career, we also bring a cohort dimension to the analysis of match quality for a number of reasons. First, female labor market participation has increased significantly over the last several decades. Second, 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 by reason of 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 (Addison, Ozturk, and Wang 2018). Third, there has occurred a shift in perceived gender roles and in the formation and nature of relationships. As a result of these developments we might expect the new generation of females to be less constrained in their searches and have better matches. To the best of our knowledge, the present paper is the first to provide comparisons across cohorts.

2 Background Literature

In analyzing the relationship between match quality and wages, we focus on occupational match quality and follow a literature that emphasizes the importance of returns to occupational tenure (Kambourov and Manovskii 2009). We will describe an occupation as a composite of the tasks to be performed and the skills and knowledge required to master them, using Occupational Information Network (O*NET) data. It has been shown that the O*NET indicators capturing these task contents of occupations and their knowledge, skill and ability requirements are highly correlated with wage outcomes (Yamaguchi 2012; Gathmann and Schönberg 2010). There is a related literature that supplements occupational task content with information on worker characteristics obtained from measures of cognitive and non-cognitive ability to establish measures of multi-dimensional skills-based mismatch (Guvenen et al. 2018; Lise and Postel-Vinay 2016). Our paper follows the methodology of the latter line of research, which reports that multi-dimensional skills-based mismatch leads to wage losses even if the underlying mismatch-generating mechanism differs in the two studies. Thus, mismatch in Lise and Postel-Vinay

(2016) arises from search frictions. The costs of mismatch are found to be considerable, especially for cognitive skills. Depending on the worker's initial skills, mismatch can lead to 8 to 22 percent less output over a career. Mismatch in Guvenen et al., on the other hand, stems from workers having imperfect knowledge of their own skills: only by moving from occupation to occupation do they update their priors and find better matches. Guvenen et al. find strongly negative wage effects of mismatch, especially if it is persistent. They report that the the earnings difference between the 90th percentile and the 10th percentile of mismatch is 4.4 percent after 5 years of occupational tenure, widening to 7.4 percent after 10 years. Cumulative mismatch yields a wage difference of 8.9 percent. Mismatch mostly slows wage growth rather than having an immediate effect on levels. It is also found that the negative effects of mismatch vary by education, being much larger for college graduates and in particular for the cumulative mismatch measure.

In their definition of mismatch, these two papers closely resemble our own treatment. This is particularly true of Guvenen et al. as regards the modeling of current and cumulative mismatch and the incorporation of learning by doing in the model through occupation tenure mismatch interactions. However, both papers only address mismatch among males and fail to consider gender as a possible component of mismatch or the return to mismatch. Crucially, then, they do not address gender wage disparities from mismatch.

Even if the new mismatch literature does not explore gender differences, this is not true of earlier studies that define mismatch from the perspective of educational gaps; specifically, discrepancies between a worker's highest completed schooling and the years of education required for a job. That said, this literature has largely reported an absence of material gender differences (see Hartog 2000; Battu 2000). Problems with this commonly described 'over-education' literature include the subjective bias of self-reporting in the absence of formal education measures, estimates that are potentially biased by the failure to consider the endogeneity between the reasons for accepting a position and wages, and its use of a simplistic unidimensional definition of skill.

However, one such study bucking the trend in reporting material gender differences also seeks to provide a more thoroughgoing attempt to distinguish between the components of mismatch. Mavromas et al. (2013) differentiate between education and skill mismatch using Australian longitudinal data. Mismatched workers in this paper consist of those who are over-educated, over-skilled or both. Wage penalties for each category of mismatch are reported for women whereas for males a wage penalty attaches only to those who are over-educated and over-skilled.[1]

In re-inserting gender into the contemporary methodology, our approach also provides an important link with Goldin (2014), who attributes most if not all of the gender gap in pay to the incentive firms have to disproportionately reward those individuals who work long hours or who work particular hours. We argue that that if other (i.e. flexible) jobs are limited to only a few occupations, then females will experience greater mismatch when flocking into them. We test whether a Goldin-type explanation has any purchase in determining match quality by constructing a measure of *occupational flexibility* which we then link to occupational mismatch.

More generally, we note that Goldin's approach is echoed in the growing literature on 'family friendly' firms and jobs (see, for example, Kleven et al. 2015, 2018; Hotz et al. 2017). However, although firm family friendliness counteracts the wage penalty attaching to

[1] A study that intermediates between the over-education literature and the multidimensional skills mismatch literature is Johansson and Katz (2007), which uses a job analysis measure of skills mismatch rather than average/modal education. Another study by Liu et al. (2016) offering an adjusted-earnings based measure of mismatch is notable for its emphasis on persistence in mismatch, a notion that is also exploited in the present study.

motherhood in this literature, observe that this offset is achieved at the price of occupational progression because family friendly firms in practice employ a less-skilled workforce, and offer lower wage dispersion and reduced scope for career advancement. Accordingly, this literature would not agree with the proposition that discrimination in temporal job flexibility suffices to close the gender wage gap as Goldin suggests is the case (see also Cardoso et al. 2016).

Finally, if the effect of motherhood and implicitly the existence of a *mommy track* is one key to our analysis, a corollary is the postponement premium (Miller 2011). That is, we examine the effect of delaying childbirth on mismatch and thence upon earnings, in effect offering a human capital explanation for the timing effect. However, we do not here directly examine the proposition that the choice of timing of childbearing is also a function of expected future earnings. Both factors and not just the former are likely to be important.

3 Data and Measurement Issues

3.1 Data Sources and Sample Construction

Our main data sets 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 to 17 years in 1997. 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, thereby excluding 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 are over 16 years of age, who were not working before the sample period, who are not full-time students, and who are currently in dependent employment in for-profit organizations and have worked for more than 1,200 hours in the preceding two years. We also exclude those who work for no pay or who report hourly wages of less than $1, as well as observations for which the wage entries are clearly in error.[2] Having also excised those with missing information on any of the variables used in the analysis and individuals without valid Armed Services Vocational Aptitude Battery (ASVAB) scores, our final sample comprises 42,022 person-year observations (with 1,890 males and 1,980 females) from the NLSY79 and 15,893 person-year observations (1,588 males and 1,434 females) from the NLSY97 over the survey periods analyzed. Table 1 reports the observation losses due to each sample inclusion criterion, while Table 2 provides basic descriptive statistics for our two NLSY samples.

[Table 1 and Table 2 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 ASVAB test scores. Such measures are unavailable in otherwise similar panel data sets. We use the age-adjusted percentile scores of respondents on the subtests of the ASVAB as the basis of our individual skill measures (see the next subsection).

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. We mapped all available NLSY79 and NLSY97 occupation codes using the guidelines developed by Dorn (2009) so as to be able to exploit the full extent of the data panel available. (Further details on occupational code mapping are given in Appendix A.)

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

3.2 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 Career Exploration Program. This 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 conducted the ASVAB tests around their inception; specifically, for the first round of the NLSY97 and the second year of the NLSY79. All NLSY79 respondents and about 80 percent of the NLSY97 sample participated in the computer adaptive test of the Armed Services Vocational Aptitude Battery (CAT-ASVAB).[3]

We consider four composite skill endowment measures: Mathematical, Verbal, Science/ Technological/Mechanical (STM), and Social. For the first three composites, for all those in the NLSY samples with valid test scores, we constructed measures using percentile ranks on select ASVAB subtests. Specifically, for the verbal skills composite 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* using the weights provided by the NLS staff.[4] We then converted these composite scores to percentile ranks, which range

[3] 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.

[4] We thank Steve McClaskie and other NLS program staff for their help in providing us with the weights and for assisting us with the program that creates the weighted composites. This program also adjusts the raw scores by age within 3-month birth cohorts.

between 0 and 1 (that is, from 0 to 100 percent, where, for example, 0.75 refers to the 75th percentile).[5]

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. (2018). 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 (retrospective) sociability at age 6 along with their rank on the Rotter Locus of Control Scale and the Rosenberg Self-Esteem Scale.[6] The NLSY97 does not ask these sociability questions nor does it collect data on 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.[7] Table 3 reports descriptive statistics on the skill endowments for each NLSY cohort by gender and educational attainment.

[5] This approach is similar to that used by Guvenen et al. (2018) other than for the inclusion of STM scores. There is no consensus in the literature as to the construction of the ability measures. Although almost all studies utilize ASVAB test scores, they select different ability dimensions or different subtests for measurement of these dimensions. Our results were robust to variation in measurement, such as the exclusion of STM skills by Guvenen et al. (2018) and the restriction of ASVAB measured abilities to *cognitive*, *manual* and *social* by Lise and Postel-Vinay (2016) who analyze mismatch by separate ability dimensions and eschew use of an aggregate measure.

[6] 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.

[7] Deming (2017a) uses two additional questions on high school participation in clubs and sports for his analysis of 1979 cohort data. For cohort differences, he switches to a two-question measure. As noted, we only use two sociability questions for the NLSY79 which is consistent with his cohort analysis. The literature displays multiple ways of measuring social skill or abilities. In Guvenen et al. (2018), for example, 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. Again, our results were robust to alternative measures using either of these definitions.

[Table 3 near here]

In our analysis, every occupation is defined by the combination of knowledge, skill, and abilities (KSAs) it requires. We use the O*NET database to determine the requirements of each occupation.[8] For each of the ASVAB subtest scores that is used as components of the first three skill endowments in Table 3, there is a corresponding knowledge, skill, or ability that is associated with a task performed or a worker quality required in that occupation. The DoD has a mapping between ASVAB subtests and knowledge, ability and skill measures in O*NET which they utilize to assign military personnel. The mapping is also used by others, such as high school counselors to recommend careers to ASVAB-participating high school students. Our match quality measure is based on the ranking comparison strategy that is used by these groups. This mapping is provided in Appendix B. However, there is no social skill component to these DoD assignments. Again following Guvenen et al. (2018) 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 with the NLSY data.

Even though the measure of skill endowments we use is superior to the unidimensional measures of the previous literature, it has the limitation that worker endowments are measured pre-labor market and do not evolve over time. For example, endowments do not change with learning by doing and professional education is not a component of the skill endowment set. Our favored interpretation would be that this skill measure is more about the potential of the worker –

[8] We are using the 2007 version of the O*NET database, after Hirsch and Manzella (2015). We are indebted to Barry Hirsch for kindly providing us with these data. Original O*NET data grouped occupations using Standard Occupational Classification (SOC) codes. Hirsch and Manzella (2015) mapped these codes to COC 2002 codes. We once again used Dorn’s (2009) mapping to link the O*NET data on occupational KSAs to individual occupations in the NLSY data.

potential to learn and potential to build the set of skills required by any given occupation – than it is about his or her contemporaneous skills. Again, it would be preferable to have formal contemporaneous endowment measures (as well as measures of occupational requirements that evolve over time). However, since we are interested in relative realizations of the matches by gender, our mismatch measure may be less vulnerable to these measurement limitations if skills do not differentially evolve by gender. In any event, we will indirectly tackle some of these measurement issues, such as learning by doing, in our estimations.

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.[9] 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_{ijc} denote individual *i*'s O*NET occupational requirements for skill *j*, in occupation *c*. The degree of skill mismatch for individual *i* for skill *j*, in occupation *c* is calculated as

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

such that the lower the value of the sum of q's over all 4 dimensions, or aggregate mismatch, the better the skills are matched. In our empirical application, this aggregate mismatch measure is our main outcome variable. In generating this measure, we used equal weights for all skills. Our

[9] We also used 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. (2018).

results were not sensitive to alternative weighting schemes. For ease of interpretation of the coefficient estimates, we rescaled this measure to have a standard deviation of one.

[Table 4 near here]

Table 4 offers a descriptive view of the aggregate mismatch measure by skill dimension, experience, and gender for each cohort. The table offers several important pointers. 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] and [4]). Third, and in consequence, 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.[10] 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 what follows, we shall mainly concentrate on the size of mismatch and the effect of mismatch on wage outcomes. However, we will also create dummy indicators to identify the direction of mismatch and test whether women are more likely to be over-qualified. We will also investigate whether or not the quantity of mismatch and wage effects of mismatch differ by skill type.

4 Conceptual Framework

The conceptual framework guiding our thinking is similar to that outlined in Guvenen et al. (2018), but containing additional layers to accommodate gender and cohort differences. In their setup, an individual worker's productivity is a positive function of match quality. As a result, all

[10] It will be recalled that the 1997 Cohort is still young, with the average worker having less than 10 years of experience.

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 as and when they can.

The first additional layer is inspired by Goldin (2014), who identifies differences in the 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, 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 after 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. *Vulgo*: 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. Although 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.

5 Econometric Analysis

Our analysis proceeds as follows. Beginning with the NLSY79 cohort, we first document the determinants of the magnitude of the total amount of mismatch in the labor market. For this 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 examine how the need for flexibility and household decision dynamics may play a role in this process. Only then do we 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 shall test whether or not, all else equal, the two cohorts display different match quality by gender in this early-career phase.

Finally, returning to the NLSY79 cohort, we analyze match quality and the wage effects of match quality by skill type. Specifically, we investigate whether mismatch differs by skill type and if there are any gender differences in this regard. The wage penalties from mismatch associated with each skill type are also examined.

5.1 Gender and Mismatch

We first consider the role of gender in mismatch in Table 5, where the dependent variable is the total amount of mismatch. The first column of the table includes only the female dummy. This specification reveals no significant gender differences in match quality. We next allow the relationship to differ by education status (but condition on nothing else); specifically, we distinguish between persons with at least a college degree and those individuals without one. It now emerges that women with less than a college degree are less mismatched while more educated women are more mismatched than their male counterparts. For the less well educated group, the match quality gender difference is about 7 percent of a standard deviation in favor of females. Women with college degree or higher educational level, on the other hand, are found to have on average about 9 percent of a standard deviation (0.0876 = 0.1588 - 0.0712) greater mismatch than men with at least a college degree. For their part, males of the higher educational attainment group are about 17 percent of a standard deviation less mismatched than their less well educated counterparts.

When, in the third column of the table, we include individual characteristics such as race and adjust for the individual's own skill endowments and the skill requirements of the occupation, two interesting things happen. First, the match quality advantage of less well educated females over males of similar educational attainment is reduced and is no longer statistically significant. Second, the match quality difference between highly educated and less well educated males deepens by about 10 percentage points to 26 percent of a standard deviation, indicating that some of the mismatch is masked if one does not factor in individual differences in human capital and occupational requirements. In our baseline specification, given in the fourth column of the table, where we also control for the labor market experience and occupational and employer level tenure of the worker in quadratic form, and their interaction with the individual's average skill scores and occupational requirements, these effects are largely unchanged. Here, we observe that females with a college or more education have on average 12 percent of a standard deviation greater skill mismatch (0.1184 = 0.1509 - 0.0325) than 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 statistically significant. To repeat an earlier point, for males having at least a college education significantly reduces occupational mismatch by a little more than 26 percent of a standard deviation. The role of education in reducing mismatch is not as important among females – at only 11 percent of a standard deviation (-0.1092 = -0.2649 + 0.1509).

[Table 5 near here]

There is one obvious issue that remains to be addressed. 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 approach this endogeneity problem by instrumenting for employer tenure, occupational tenure, and total experience. Following Altonji and Shakotko (1987) and Guvenen et al. (2018), we instrument individuals' employer tenure and occupational tenure variables with their deviations from spell-specific means, controlling for the possibility of multiple spells of employment with the same employer or in the same occupation.[11] Total labor market experience is instrumented in a similar fashion.

[11] $IV_{emp} = T_{emp} - \overline{T_{emp}}$ 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 remains in the labor market and $\overline{T_{exp}} = \frac{1}{S}\sum_{t=1}^{S} T_{exp_{i,j,t}}$, S being the total number of spells that an individual is observed to be in the labor market. For example, if an employer has a 3-year spell with an employer, in

The fifth column in Table 5 reports results for this IV specification. For the estimates that are of principal interest we observe no significant change in magnitude or sign vis-à-vis the baseline specification.[12]

Another concern is the role of part-time jobs in influencing match quality. Our sample restrictions were designed to lead to the inclusion of only those individuals with sufficiently strong labor market attachment. However, as our 1,200 hours a year restriction roughly translates into 25 hours a week, the sample will certainly include some part-timers (about 15 percent). The last two IV model estimates in Table 5 seek to discover whether part-timers are different from full-timers. It can be seen that gender differences in match quality are less pronounced for part-timers than for full-timers. For example, female college graduates who are employed part-time are not significantly more mismatched than their male counterparts, which is in sharp contrast with the results for full-timers. Females may be less likely to remain employed full time after life events such as birth of a child, but there are no obvious mismatch implications of this particular change in status. This may be because better matched college graduate females may be those who stay in their jobs with reduced hours while their worse matched counterparts cannot and eventually drop out. Yet our part-timers may not be representative of all part-timers given their likely stronger labor market attachment. Accordingly, even if it is abundantly clear that part-timers are not driving our results, this recognition does not eliminate the need for a separate and specific examination of part-time work and mismatch in future work.

---

the first year the value of the instrument will be -1, in the second year it will be 0, and in the third year it will be 1, because 2 is the average experience over these three years of observations ((1+2+3)/3).

[12] This instrumentation strategy, although commonly used, is not an ideal one. It relies heavily on the assumption that fixed components of the error terms can be linearly decomposed and that there is no correlation between the fixed effects corresponding to different tenure and experience variables. In simulation exercises, Pavan (2011) has shown that such IV estimates can be strongly biased when these assumptions fail. He also points out that the bias is almost non-existent for workers with at least a college education. Thus, we incline to view that our estimates for the college educated are unaffected. Given the bias documented for the non-college educated, however, we will refrain from overstating insignificant results for this group. Our maintained hypothesis is that these endogeneity issue are likely gender neutral and do not affect our conclusions regarding mismatch and gender.

5.2 The Role of Household Formation and Fertility

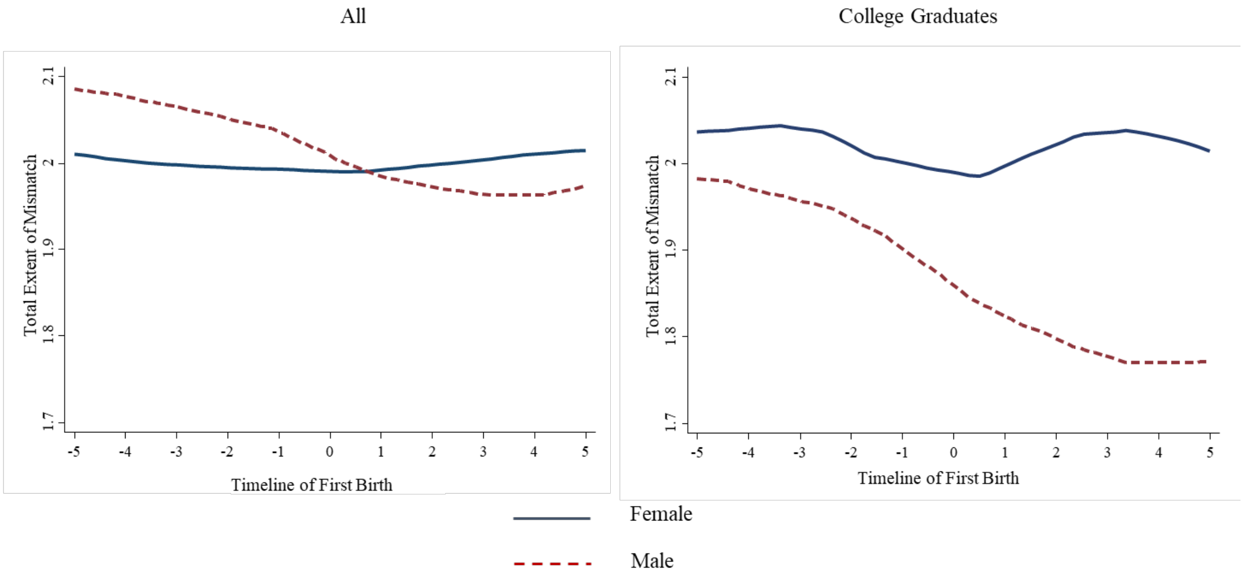

Figure 1. Fertility and Mismatch over the First Birth Timeline

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. The descriptive association between fertility and match quality is graphed in Figure 1 for the NLSY79 cohort. The figure maps match quality (as described in the preceding section) first for all males and all females, and then for college graduates only, through the timeline of the first birth. 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 worker mismatch in both panels of the figure. By contrast, over the same timeline, male worker matches actually improve, especially among those with a college degree. Now this descriptive look at the fertility-mismatch relation is likely biased. Fertility, marriage and mismatch maybe endogenously determined while there is selection in who marries and who chooses to have a child. Controlling for possible confounding factors and addressing these endogeneity issues, Table 6 offers an initial look at the effect of marriage and fertility on the size of mismatch. Three empirical models are presented. The first addresses the issue of endogeneity in respect of the tenure and experience variables mentioned earlier (see the pooled IV specification in Table 5). The second exploits the panel nature of our data and addresses the possibility of unobserved individual heterogeneity within a fixed effects setup. Here we rely on within-variation to identify the effect of marriage and children on occupational match quality, specifically changes in marital status and the birth of the first child. The third and final model combines efforts to alleviate both concerns in reporting estimates from 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 about 7 percent (0.0664 = 0.0971 - 0.0307) of a standard deviation relative to childless females. The effect of simply having a child is much smaller, and is in fact not statistically significant for males (-0.0307). For its part, marriage is not significantly related to mismatch. In the fixed effects specifications, the effect of fertility is more pronounced for females. Women who have at least one child are about 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. Gender specific regressions reaffirm the role of children in increasing mismatch for females (10 to 11 percent of a standard deviation in the last two specifications) while the effect is still negative and non-significant for males.

[Table 7 near here]

Our descriptive statistics provided 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 is arguably 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 1.95; that is, females with children are almost twice as likely to be over-qualified. 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 have shown 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, we next consider in Table 8 the change in mismatch along the fertility timeline, again for the NLSY79 cohort. For an interval extending up to six years after the first birth, females make worse matches than they did within 3 years prior to that birth event. The disparity is approximately 6 percent of a standard deviation in the OLS results, 8 percent of a standard deviation for the IV-FE regression,[13] and up to 10 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

[13] Here we are only instrumenting for the tenure and experience variables, and utilize the FE setup to access the issues of endogeneity and selection in fertility.

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 in mismatch shortly following the first birth may accumulate, leading to even greater mismatch later in a woman's career. It is therefore of no small interest to explore whether a delay in fertility might reduce the size of mismatch induced 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.[14]

In a second set of regressions (IV-2) we also instrument 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 setup. When comparing the IV results with the OLS estimates, we can conclude that the age at first birth is endogenous with respect to the degree of occupational mismatch. Overall, the OLS estimates are biased upwards; that is, ignoring endogeneity clearly underestimates the gains in match quality resulting

[14] This is not one of the instruments commonly used in the literature but is the only one that is available for both men and women in our data set. It has been documented that siblings' fertility behavior is highly correlated (Lyngstrad and Prskawetz 2010). More specifically, it has been suggested that shared upbringing, close social interactions with siblings, and peer effects result in a positive relationship between the individual's own birth timing and that of his or her sibling. This instrument passes all the tests to be a valid instrument. We present the first stage results in Table D.1. These indicate that age at first birth variables are only endogenous for females. Thus, we present IV estimates only for females. Estimates using miscarriages as an instrument for motherhood timing after Miller (2011) are also presented in Tables D.2 and D.3. These estimates imply even stronger results than do ours. However, this variable is only available for women and for the 1979 sample alone.

from a delay in childbirth – at least for some women. It might also be the case there is selection into delayed fertility by individuals who would otherwise suffer the biggest losses. Unfortunately, our empirical strategy cannot decide this important issue.

5.3 Occupational Flexibility and Mismatch

We next 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 measures: *time pressure, contact with others*, *establishing and maintaining interpersonal relationships, structured vs. unstructured work*, and *freedom to make decisions*.[15] The higher the score, the more flexible is the occupation. In Table 10, we see that among those persons with at least one child, a one standard deviation more flexible job is associated with a (roughly) 4 percent of a standard deviation higher degree of mismatch for males and a 6 percent of a standard deviation higher degree of mismatch for females.

[Table 10 near here]

To illustrate the relationship between occupational flexibility and mismatch over the first birth timeline in Figure 2 we plot the predicted amount of mismatch at different points on the fertility timeline for first-time parents in the highly flexible and highly inflexible occupations (respectively, the 90th and 10th percentile flexibility values as observed in the data). The figure

[15] 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; (iii) *Establishing and maintaining interpersonal relationships: 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 of 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 rescaled this score to have a standard deviation of 1.

shows how females gravitate to work in occupations offering high flexibility at the expense of a better skill match after childbirth. No such pattern is observed for males. The estimates used in the figure to capture the relationship between mismatch and occupational flexibility over the first birth timeline are provided in Table D.1.

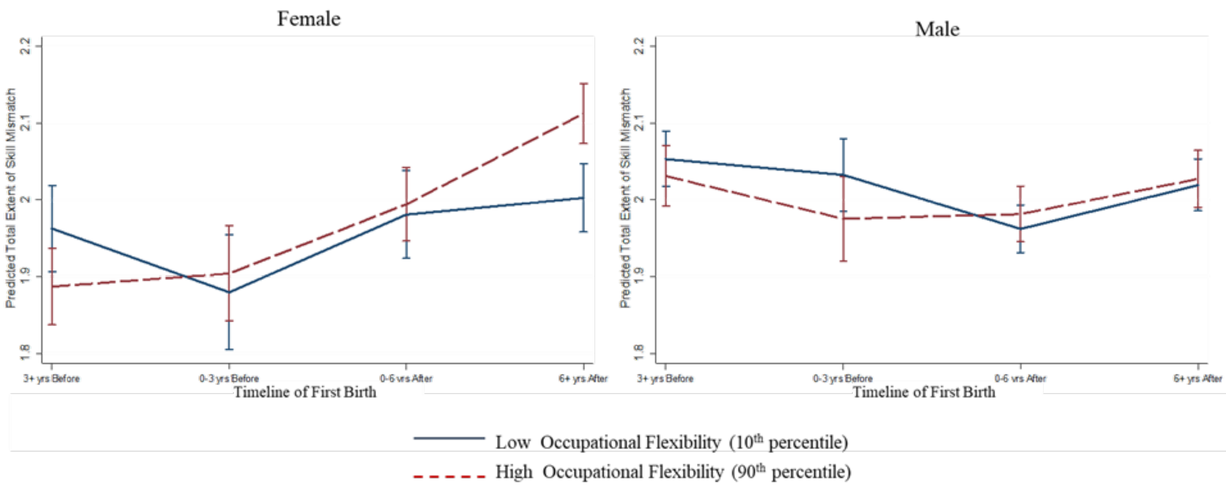

Figure 2. Flexibility and Mismatch over the First Birth Timeline

5.4 Wages and Mismatch

Thus far we have shown that, contrary to the prediction of search models, women's match quality does not improve over the course of their careers. Women, and especially the most educated among them, are more mismatched than their male counterparts. Following Guvenen et al. (2018), but now for both genders, we chart in Table 11 the wage loss associated with this mismatch. Also as in Guvenen et al., we model the wage loss associated with the total extent of mismatch along three dimensions: (a) a threshold penalty, (b) a decreasing return over the career path, and (c) a wage penalty associated with cumulative past mismatch.[16]

In Table 11 outcome variable is log real hourly wage (in 2002 dollars). In the first

[16] We calculate cumulative mismatch in the same manner as do Guvenen et al. (2018); 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 refer to it as cumulative mismatch.

column of this table we only include the measure of mismatch in the current occupation.[17] Our results imply for males that an individual who is one standard deviation worse matched than the average amount of mismatch will earn 5 percent less (-0.0236*2) than an individual who is one standard deviation better matched than the average. The wage effect is a little less for females at 4.4 percent ([-0.0236 + 0.0014] * 2), although 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 these added interaction terms are not significant. However, when the cumulative mismatch – that is, the measured history of mismatch, as described above – is added to the model (the third column of results) the effect of current mismatch becomes insignificant. Cumulative mismatch, on the other hand, implies a 12 percent lower (-0.0598*2) wage for males who have a one standard deviation worse than average match history compared with those with a one standard deviation better than average match history all else constant. Cumulative mismatch in this model does not have differential effects on female wages.

[Table 11 near here]

Next, in the fourth column of the table, 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, on average, college graduates suffer much stronger current and cumulative mismatch wage penalties than their less well educated counterparts. A one standard deviation higher than average current mismatch implies a higher wage penalty for college graduates of

[17] Here we are only reporting the results for the instrumental variables specifications used subsequently 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.

about 3 percent for males and 4.8 percent for females on average. And cumulative mismatch has a somewhat stronger wage penalty for males (5.1 percent) but not for females (4 percent), again relative to non-college graduates.

Similar differences in wage penalties are reported by educational levels for the gender-specific regressions in the last two columns of Table 11. Even though current mismatch does not imply a significant wage penalty for either males or females among non-college graduates, for college graduates of both genders a one standard deviation greater mismatch implies about 4 percent lower wages (-0.0362 in the case of males and -0.0429 for females) on average during the first year in an occupation. Although female college graduates experience lower wage losses then their non-college educated counterparts who are equally mismatched over their tenure (a coefficient of 0.0029), no such significant differences are observed among males. For its part, cumulative mismatch is more punishing for males and females with college degrees relative to their non-college educated counterparts, albeit only significantly so in the case of males.

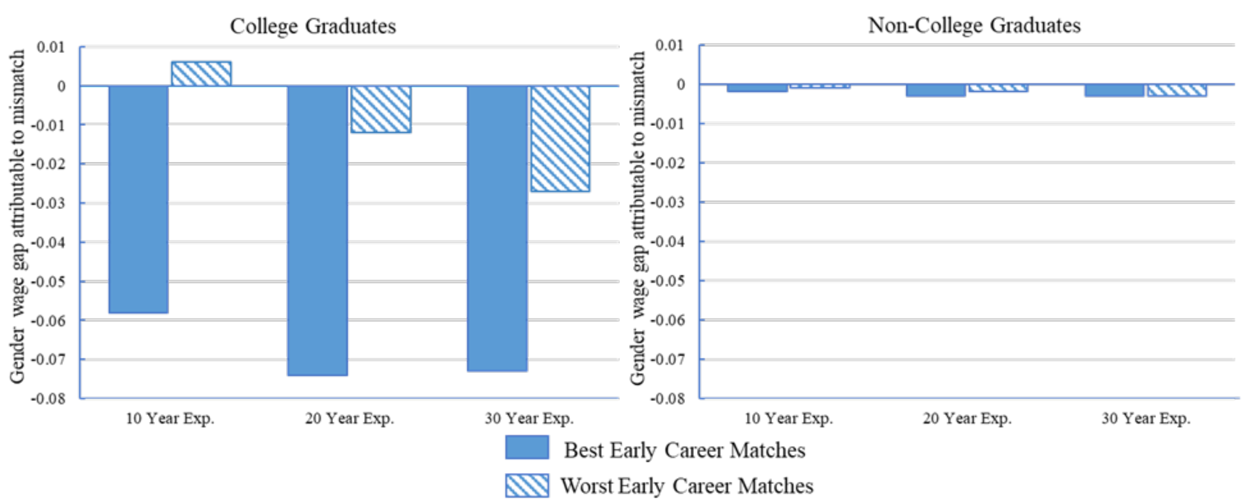

Figure 3. Mismatch and the Gender Wage Gap over a Career by Education

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 there are potentially more children in the household. We also demonstrated that cumulative mismatch has a more detrimental effect than current mismatch. Although there are no significant gender differences in the 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. To illustrate the magnitude and persistence of wage effects of gender disparities in match quality we conduct a simple exercise. In Figure 3 we link the gender difference in mismatch and wage penalties to both current mismatch and cumulative mismatch. Distinguishing here between college educated and non-college educated individuals, we chart the progression of gender wage disparities over different stages of a career at 10 year intervals (vis. 10 years, 20 years and 30 years of experience) that can be attributed to match quality alone. We do so separately for those individuals with the best and worst early career matches.[18] We see that gender disparities due to mismatch are most pronounced for college graduates with the best early career matches. Female college graduates with the best early career matches earn approximately 7 percent less than their male counterparts after 30 years with average current and cumulative mismatch experience. The gap due to mismatch is little over 2 percent after 30 years among the worst early career matches in this group. On the other hand, irrespective of the quality of early career matches, gender wage gaps are muted for those without a college education.

Given our earlier findings regarding the role of fertility delay in mismatch experience, we can also visualize the wage penalty that can be attributed to motherhood for different schooling levels and by different ages at the first birth. Figure 4 first graphs the implied wage differences at

[18] Early match quality is determined according to an individual's mismatch over the first 5 years of experience. A distinction is drawn between the top (worst matched) and bottom (best matched) deciles of the mismatch distribution. Current and past mismatch values are averages for the groups with the given levels of experience and early career match quality, assuming this experience is also equal to occupational tenure. From Table 11 we see that for both educational groups wage growth is not materially influenced by the degree of mismatch (captured by the coefficient of the occupation tenure and mismatch interaction). Disparities between occupational tenure and experience will not affect Figure 3 as calculations for this figure use only the significant coefficients. Calculations used in assembling Figure 3 are presented in Table D.2.

different points on fertility timeline for college graduate females who have given birth to their first child at age 25 and at age 30 relative to their otherwise identical counterparts who have not given birth. It then repeats the same exercise for females with less than a college degree. According to this figure, the motherhood wage penalty due to mismatch is higher among the college graduates, and especially those who have given birth early. Specifically, a female college

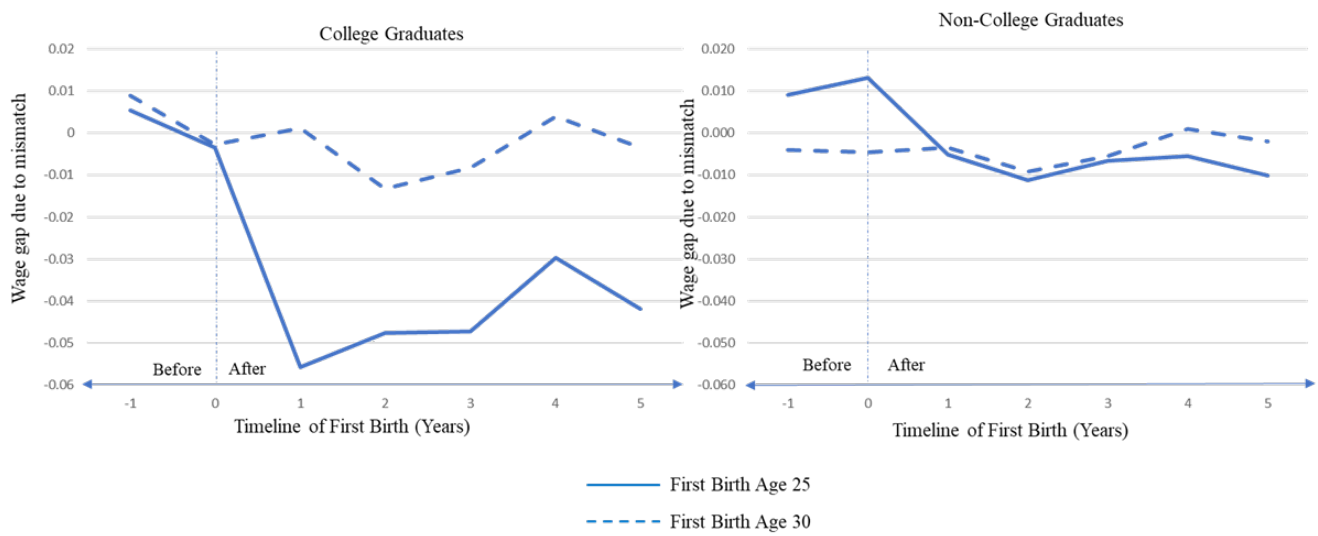

Figure 4. Mismatch and the Motherhood Wage Penalty by Age at the First Birth and Education

graduate who gave birth to her first child at age 25 would earn about 4 percent less by age 30 – 5 years after her first birth – than her counterpart who gave first birth at age 30 at the same point on fertility timeline (at age 35).

5.5 Cohort Differences

For this next part of our analysis we use only those NLSY79 observations for respondents when they were aged 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 used both by itself and in interaction with other variables of interest. Table 12 examines cohort differences in mismatch and over-qualification. For the ‘younger cohort’ (NLSY97), we see that college-educated females continue to be more mismatched and more likely to be over-qualified than their male counterparts (0.0328 of a standard deviation more mismatched and with about one percentage point higher probability of being over-qualified),[19] although in neither case significantly so. On the other hand, among those without college degrees, females are less mismatched (almost 4 percent of a standard deviation in the case of the NLSY79 and more than 7 percent of a standard deviation for the NLSY97). They are also less likely to be over-qualified than their male counterparts of both generations: about 1.4 percentage points less in the case of the NLSY79 and about 2.5 percentage points less for the NLSY97. Finally, Table 12 suggests that among both college graduates and non-college graduates the younger cohort of both genders perform worse, recording greater skill mismatch and a higher incidence of over-qualification on average than their older counterparts.

[Table 12 near here]

Table 13 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 (5.64 percent of a standard deviation more), while the opposite is true for women of that cohort (8.03 percent of a standard deviation less). Relative to the older cohort of males, they are more mismatched within 6 years following the birth of the first child when compared to the baseline interval of up to 3 years before birth (5.19 percent of a standard deviation), although this large effect is imprecisely estimated. This pattern is in strong contrast to what was observed for older cohorts, where males became better matched after having had a child, especially when the children were young (about 4 percent of a standard deviation in Table 8, and roughly 2 percent of a standard deviation, albeit statistically insignificant, here in the

[19] See Appendix E for a comprehensive comparison of predicted probabilities.

Table 13 regression). This shift implies that male millennials have assumed more family responsibilities than did the male baby-boomers before them. For younger-cohort females with at least one child, mismatch is less than that reported for their counterparts in the older-cohort (about 8 percent of a standard deviation less). With respect to the fertility timeline, we observe that younger-cohort females are also less mismatched more than 6 years after the first birth than the corresponding group of older-cohort females, although the coefficient estimate is statistically insignificant.

[Table 13 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 recent BLS (2017) report, the share of women earning more than their husbands has increased by more than 60 percent over the last several decades. Table 14 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 the table we examine whether being a breadwinner decreases the level of mismatch in households with and without children. Comparing the IV and IV-FE results, we first see that some of the breadwinner status effects are explained away once we use within-variation for identification. For example, in the case of males without a child, being a breadwinner implies a 10 percent of a standard deviation reduction in the amount of mismatch using the IV specification. Once we consider individual fixed factors however, not only is the magnitude of this effect greatly reduced (from 0.1008 to 0.0104), but also it is no longer statistically significant. For the fixed effects-only models, it is apparent that not only are female breadwinners less mismatched than their non-breadwinner counterparts but also that the breadwinner effect is modestly stronger for the NLSY97 cohort of females. For the NLSY79 male cohort there is some indication that being a breadwinner with at least one child may reduce mismatch, but this is nowhere the case for the younger cohort of males.

[Table 14 near here]

5.6 Mismatch by Skill

The most recent literature suggests that the effects of mismatch might differ by the type of skill for which the worker is mismatched (Guvenen 2018; Lise and Postel-Vinay 2016). In particular, most studies suggest the wage penalties appear highest for cognitive skills. Workers can make up for the manual and routine skills that they lack, but cognitive skills when under-qualified are harder to rectify. Also, when over-qualified, unrealized returns to cognitive skills are higher than for all other skill dimensions. Thus, one pressing question for our own inquiry is whether gender differences in mismatch vary by skill type; another is whether the wage effects of mismatch also differentiated by skill.

[Table 15 near here]

Using IV-FE models, the upper panel of Table 15 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 4.98 percent of a standard deviation more mismatched in terms of Math skills than their non-college educated male counterparts when they have a child. There is an even larger gender difference in the extent of mismatch for STM skills in the presence of children for non-college graduates; compared with their male counterparts with at least one child, these females are 8.94 percent of a standard deviation more mismatched. This particular gender effect is not as strong in the case of college-educated individuals and the coefficient estimate is statistically insignificant (although its magnitude is large at 6.21 percent of a standard deviation). For Math skills, however, gender differences are significant: among college graduates with at least one child, females are about 11 percent of a standard deviation more mismatched than their male counterparts.

Turning to the gender-specific models, we see that non-college educated women who have at least one child are also significantly more mismatched across all skill dimensions compared with their childless counterparts by about 8 to 11 percent of a standard deviation. For their part, non-college educated males with at least one child are significantly more mismatched compared with their childless counterparts in Math skills, while college educated males are significantly less mismatched throughout (other than in the case of Social skills).

The results for over-qualification, shown in the lower panel of the table, are frankly mixed. In the important category of Math skills, however, college-educated women are also more likely to be over-qualified when compared with both their male counterparts and childless women.

[Table 16 near here]

Finally, Table 16 seeks 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 for Math skills for males disappears. However, for females on average we observe a negative and significant threshold mismatch effect for STM skills, although this disappears after about 5 years of tenure. 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 cumulative mismatch in terms of Math skills has the largest wage effects for males. For females, on the other hand, penalties are lower for Math cumulative mismatch but are of the same order of magnitude for cumulative mismatch on Verbal and Social skills.

6 Conclusions

This study has examined match quality by gender and different cohorts using the NLSY79 and the NLSY97. It extends the empirical methodology of Guvenen et al. (2018) in which skill mismatch is based on the discrepancy between the portfolio of skills required by an occupation and the portfolio of abilities possessed by the worker for learning those skills. Worker abilities or skill endowments are captured by ASVAB scores, occupational requirements are documented in O*NET database, and these two dimensions are linked using a DoD mapping.

Although our methodology improves on past mismatch arguments based on average years of schooling undertaken and the schooling required, or the use of self-reported indicators of mismatch, a number of limitations attach to our mismatch construct. First, skill endowments are fixed, having a basis in a test administered in adolescence. Second, there are unmeasured changes in occupational requirements over time. Ideally we would have a more contemporaneous measure of endowments and skill requirements. However, although skill endowment is formally fixed, a major focus of our study is to track how life events in the form of marriage and child birth affect the skill match quality for workers. Also, it will be recalled that in the regressions for match quality and for wages we include an interaction term between individuals' skill endowment and their occupational tenure to capture potential on-the-job learning.

However, the latter observation raises the vexed question of the exogeneity of tenure as the tenure variable is likely to be correlated with individual and match specific factors. In seeking to tackle this issue, we follow Guvenen et al. (2018) in using an instrumental variables strategy pioneered by Altonji and Shakotko (1987). The instrument chosen – the deviation from individuals' average length of tenure in that occupation – is the best we can do given the nature of our data set and should satisfy the exogenous variation condition. It has been criticized by Pavan (2011), who argues that (within the framework of a two-stage search model) IV estimates of this type underestimate the effect of firm-specific job matches. Even if there is no obvious way of adopting his particular estimation strategy with the data at our disposal, Pavan's reservations do have to be borne in mind in interpreting our results.

Some imprecision also attaches to our modeling of the interaction of fertility and mismatch. In recognition that the respondent's match quality might influence the timing of the birth rather than the other way round, we also instrumented the timing of the first birth. As an instrument we used the age of the respondent's sibling at first birth. This instrument has the advantage of being available for both male and female respondents in our data unlike alternatives from the literature. It has been documented that siblings' fertility behavior is highly correlated (Lyngstrad and Prskawetz 2010). More specifically, it has been suggested that shared upbringing, close social interactions with siblings, and peer effects result in a positive relationship between the individual's own birth timing and that of his or her sibling. We find in our first stage regressions evidence supportive of this contention. Thus, our instrument satisfies the inclusion restriction. Necessarily more problematic is of course negotiating the exclusion restriction, although concerns that our instrument is correlated with the outcome of mismatch or any other unobserved determinant of that outcome are mitigated by the detailed controls in our regressions.

We also investigated other possible instruments. Two alternatives instruments are “twin births” and “miscarriages,” We eschewed use of the former as an instrument (on which, see Öberg 2017), given that twin births will not capture the match quality change for workers giving birth to the first child, which event is a major concern of the present study. And although miscarriages are by no means an ideal instrument – for example, Fletcher and Wolfe (2009) argue that they may not be random events and are likely correlated with unobserved community-level factors and are only available for females – we decided to experiment with this fertility shock variable. We reported that the results of this procedure were qualitatively similar to those obtained using age of the respondent’s sibling at first birth.

Despite our various attempts to tackle biases such as endogeneity, selection, and unobserved heterogeneity, our methodology should be viewed as providing a detailed descriptive approach rather than establishing definitive causal associations. Our intent has been to provide a wide-ranging empirical discussion of gender and mismatch over a career within a framework that tells us something more about the mechanisms through which gender gaps emerge and lays the ground for the adoption of a more causal approach in subsequent research.

Our key results may be adumbrated as follows. First, females are more mismatched than males, and this result is driven by the greater mismatch among highly-educated females. Second, females with at least one child evince greater mismatch relative to their male counterparts and to females with no children. Third, marriage per se is unrelated to mismatch. Fourth, there is a degree of persistence in mismatch for females. Their mismatch persists for up to six years after the first birth, and even increases thereafter. For males, on the other hand, mismatch declines within the former interval, with no significant change thereafter. Fifth, delaying the time of childbirth reduces mismatch for females. Sixth, mismatch is greater for those individuals with children and in more flexible occupations. Although this is true for both genders, the tendency is altogether stronger for females. Seventh, mismatch results in lower earnings for both genders. The wage loss comprises three elements: current mismatch, a reduced return to tenure, and cumulative past mismatch (strictly, the weighted average of past mismatches). The biggest penalties accrue to accumulated mismatches and among those who are college educated. Eighth, although there is no significant gender difference in the effect of mismatch on wages, greater current and accumulated mismatch do contribute to increasing wage disparities between females and males over a career. Specifically, between 1 percentage point and 7 percentage points of the gender wage gap stem from current and accumulated mismatch, with the higher (lower) value corresponding to those with the best (worst) early career matches. Ninth, there is evidence of greater mismatch in mathematical skills for females, especially among college graduates with at least one child, who are also more likely to be over-qualified than their male counterparts in this dimension. However, males seem to experience higher wage penalty for mismatch in mathematical skills. That said, the effects of mismatch on wages in the cases of other skills are mostly muted for both genders. Lastly**,** younger cohorts of females appear to be doing better in terms of mismatch than their precursors although the reverse is the case for males; these cohort results offering some support to Goldin's notion of *grand convergence* and Frank's theory of *differential over-qualification*.

Some broader themes of our research findings might also be noted. Our results on skill mismatch by skill type might well suggest that, in certain fields, gender segregation and discrimination continue to apply. Another explanation for this disproportionate female mismatch is undoubtedly the division of labor in the household and traditional gender roles. 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. This is because jobs offering flexibility are not distributed evenly across the occupational spectrum. The end result is the under-utilization of a material share of labor market participants.

REFERENCES

Addison, John T., Orgul D. Ozturk, and Si Wang. 2018. The occupational feminization of wages. *Industrial and Labor Relations Review* 71(1): 208-241.

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.

Burdett, Kenneth. 1978. A theory of employee job search and quit rates. *American Economic Review* 68(1): 212-220.

Cardoso, Ana Rute, Paulo Guimarães, and Pedro Portugal. 2016. What drives the gender gap? A look at the role of firm and job-title heterogeneity. *Oxford Economic Papers* 68(2): 506-524.

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, https://doi.org/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.

Fletcher, Jason M., and Barbara L. Wolfe. 2009. Education and labor market consequences of teenage childbearing – evidence using the timing of pregnancy and community fixed effects. *Journal of Human Resources* 44(2): 303-325.

Frank, Robert. 1978. Why women earn less: The theory and estimation of differential over-qualification. *American Economic Review* 68(3): 360-373.

Gathmann, Christina, and Uta Schönberg. 2010. How general is human capital? A task-based approach. *Journal of Labor Economics* 28(1): 1-49.

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. 2018. Multidimensional skill mismatch. NBER Working Paper No. 21376. Cambridge, MA: National Bureau of Economic Research.

Hartog, Joop. 2000. Over-education and earnings: Where are we, where should we go? *Economics of Education Review* 19(2): 149-159.

Hirsch, Barry T. and Edward J. Schumacher. 2012. Underpaid or overpaid: Wage analysis for nurses using job and worker attributes. *Southern Economic Journal* 78(4): 1096-1119.

Hirsch, Barry T. and Julia Manzella. 2015. Who cares – and does it matter? Measuring wage penalties for caring work. In *Research in Labor Economics: Gender Convergence in the Labor Market* 41: 213 - 275

Hotz, V. Joseph, Per Johansson, and Arizo Karimi. 2017. Parenthood, family friendly firms, and the gender gaps in early work careers. NBER Working Paper No. 24173. Cambridge, MA: National Bureau of Economic Research.

Johansson, Mats, and Katarina Katz. 2007. Wage differences between women and men in Sweden – the impact of skill mismatch. Working Paper 2007: 13. Uppsala: Institute for Labor Market Policy Evaluation/IFAU.

Jovanovic, Boyan. 1979. Job matching and the theory of turnover. *Journal of Political Economy* 87(5): 972-990.

Kambourov, Gueorgui, and Iourii Manovskii. 2009. Occupational specificity of human capital. *International Economic Review* 50(1): 63-115.

Kleven, Henrik Jacobsen, Camille Landais, and Jakob Egholt Søgaard. 2015. Parenthood and the gender gap: Evidence from Denmark. Unpublished Paper, London School of Economics. Available at: https://eml.berkeley.edu/~webfac/auerbach/Landais2015.pdf.

Kleven, Henrik Jacobsen, Camille Landais, and Jakob Egholt Søgaard. 2018. "Children and gender inequality: Evidence from Denmark." NBER Working Paper No. 24219. Cambridge, MA: National Bureau of Economic Research.

Light, Audrey, and Manuelita Ureta. 1992. Panel estimates of male and female job turnover behavior: Can female nonquitters be identified? *Journal of Labor Economics* 10(2): 156-181.

Lise, Jeremy, and Fabien Postel-Vinay. 2016. Multidimensional skills, sorting, and human capital accumulation. Unpublished Paper, University College London. Available at: https://econ.ucalgary.ca/ sites/econ.ucalgary.ca/files/lise_postel-vinay.pdf.

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.

Lyngstad, Torkild Hovde, and Alexia Prskawetz. 2010. Do siblings' fertility decisions influence each other? *Demography* 47(4): 923-934.

Loprest, Pamela J. 1992. Gender differences in wage growth and job mobility. *American Economic Review* 82(2): 526-532.

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.

Miller, Amalia. 2011. The Effects of Motherhood Timing on Career Path. *Journal of Population Economics* 24(3): 1071–1100.

Öberg, Stefan. 2017. An introduction to using twin births as instrumental variables for sibship size. Göteborg papers in Economic History No. 22. Gothenburg: Department of Economy and Society, School of Business, Economics and Law. Available at: https://ideas.repec.org/p/hhs/gunhis/ 0022.html.

Pavan, Ronni. 2011. Career choice and wage growth. *Journal of Labor Economics* 29(3): 549-587.

Simpson, Wayne. 1990. Starting even? Job mobility and the wage gap between young single males and females. *Applied Economics* 22(6): 723-737.

Summerfield, Fraser. 2016. Matching skill and tasks: Cyclical fluctuations in the over-qualification of new hires. Unpublished Paper, University of Aberdeen. Available at SSRN: https://ssrn.com/ abstract=2754027 or http://dx.doi.org/10.2139/ssrn.2754027.

Topel, Robert H., and Michael P. Ward. 1992. Job mobility and the careers of young men. *Quarterly Journal of Economics* 107(2): 439-479.

Yamaguchi, Shintaro. 2012. Tasks and heterogeneous human capital. *Journal of Labor Economics* 30(1): 1-53.

Appendix A: Occupational Codes in the NLSY Surveys and Mapping across Different Coding Systems

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 Census codes, in addition to the 1970 codes, until 2002. After that year, the 2002 COC was used to code occupations, and subsequent to the 2010 round the 2010 COCs were also provided. For its part, even though the first five rounds of the NLSY97 employed 1990 codes to classify occupations, the 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 were described by their 3-digit 1980 CIC in the NLSY79 until 2000; thereafter, 4-digit 2002 CICs are used. 2002 CICs are available for all rounds of the NLSY97.

In order to map all occupation codes across survey years, the 2002 Census Occupation Codes (COC) are first converted to 2000 COCs 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 September 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 Planning 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 the do-files downloaded from the same Dorn website. The 6 groups, which are also used by Autor and Dorn (2013), are as follows: managerial and professional specialty; technical, sales, and administrative support; services; farming, forestry, and fishing; precision production, craft, and repair; and operators, fabricators, and laborers. There are 14 industry and sector groups. All public employees are assigned a single 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 B: The Mapping of ASVAB and O*NET Components

There are 10 ASVAB subtests (the 1997 version [CAT-ASVAB] has 12). Seven among these are grouped into three composites: Verbal, Math, and Science and Mechanical.[20] We follow DoD guidelines in mapping ASVAB subtests to these composites and then to O*NET occupational knowledge, skill and ability components (KSAs). There are 110 KSAs – knowledge (sets of facts and principles needed to address problems and issues that are part of a job), skills (the ability to perform a task well), and abilities (enduring talent that can help a person do a job) – subsets of which are required to perform successfully in each occupation (O*NET 2010). For each of the occupations in the O*NET database, either expert job analysts, job supervisors, or job incumbents rate the degree of importance each of the KSAs and the degree of proficiency needed in each for satisfactory performance in that particular occupation (ASVAB Technical Chapter accessed at https://www.asvabprogram.com/pdf/ASVAB_CEP_Technical_Chapter.pdf). KSAs capture what workers in an occupation are expected to do, not what current workers in an occupation are doing or are capable of doing, although they are highly indicative of these average worker characteristics. In this way, KSAs are analogous to the item content of a test, as both perform the same function of operationalizing the domain in question. Therefore, linking ASVAB test content and scores with O*NET occupational requirements is a natural next step and is achieved through an analysis of the relationship between ASVAB subtests and the KSAs that best describe particular occupations. For this purpose, as a first stage, two experts identify 26 of the 110 KSAs from O*NET as possibly related to the ASVAB Verbal, Math, and Science and Mechanical/Technical Composites. Next, a larger group of expert judges (9 for verbal and math and 14 for STM), comprising industrial/organizational psychologists, other types of psychologists, and psychometricians, score the relatability of each of these 26 individual KSAs

[20] *Shop Information, Auto Information,* and *Assembling Objects* are not used for 1979 and in addition to these *Coding Speed* and *Numerical Operations* are not employed in 1997.

and particular ASVAB subtest. A high correlation across judges' scores is required to establish the links. The linkage shown below is drawn up in light of this factor analysis:

ASVAB COMPONENT	O*NET Knowledge/Skill/Ability	O*NET COMPONENT
Verbal		
Word Knowledge	Ability	Inductive Reasoning
Paragraph	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

More information on knowledge, skill and abilities can be found at https://www.onet center.org/dictionary/22.3/excel/. More information on ASVAB subtests can be found at https://www.bls.gov/nls/nlsasv79.htm and a sample ASVAB score card can be found at https://www.nlsinfo.org/sites/nlsinfo.org/files/attachments/170816/ASVAB Score Report Sample.pdf.

Appendix C: Instrumental Variables for Fertility Timing

Siblings' Average Age at the First Birth

In the light of the possible endogeneity between occupational match quality and an individual's fertility timing, Table 9 instruments individuals' relative/absolute ages at first birth using their siblings' average age at the birth of their first child. The first stage results for IV-1 estimations are presented in Table C.1. As can be seen, individuals' fertility time is strongly correlated with their siblings' fertility time. A one-year increase in siblings' average birth age will result in a 0.15 to 0.2-year (about 2 to 2.5 months) increase in an individual's absolute/relative age at first birth.

The table also reports the tests of relevance and validity of the instrument. Test statistics indicate that the age at first birth variables are endogenous for females only, and our instrument is significantly and positively correlated with our endogenous variable.

Table C.1. First Stage Regression for Siblings' Average Age at First Birth as Instrument

	Male		Female	
	Absolute Age	Relative Age	Absolute Age	Relative Age
Siblings' average age at first birth	0.2272	0.2264	0.1629	0.1497
	[0.0401]**	[0.0409]**	[0.0437]**	[0.0480]**
Control for demographic variables	YES	YES	YES	YES
Control for tenure and experience variables	YES	YES	YES	YES
R-squared	0.31	0.28	0.33	0.31
Tests of endogeneity of birth-age or relative birth-age				
Durbin-Wu-Hausman chi-sq test (p-value)	0.51	0.58	0.03	0.03
Underidentification test (Kleibergen-Paap rk LM statistic):	29.61	28.25	13.44	9.65
Weak identification test (Cragg-Donald Wald F statistic):	437.83	401.93	187.98	126.12

Notes: This table presents the first stage results and the relevant IV validity tests for the IV-1 estimations in Table 9. The dependent variables are individuals' relative/absolute ages at first birth. ** denotes significance at 0.01 level.

Alternative Instrumental Variable: Miscarriage at the First Pregnancy

Our second instrument is an unconventional one. The literature on the wage penalty of motherhood instruments fertility timing in other ways for females. We will test the robustness of our results to choice of instrument using miscarriage at the first pregnancy to instrument the age at first birth after Miller (2011). As the NLSY79 data only contain information about a woman's first pregnancy outcome, we conduct the estimations for females alone. In constructing an instrument for the age of first birth of women whose first pregnancy ended in an abortion, we use the timing of the first non-aborted pregnancy.[21] The first stage regression results are presented in Table C.2. As shown in the table, the experience of a miscarriage at the first pregnancy leads to a 2-year fertility delay. This instrument is both relevant and valid as indicated by the tests statistics reported below.

Table C.2. First Stage Results for the Alternative Instrumental Variable: Miscarriage at the First Pregnancy

	Female	
	Absolute Birth Age	Relative Birth Age
Miscarriage at the first pregnancy	1.991 [0.153]**	2.0933 [0.175]**
Control for demographic variables	YES	YES
Control for tenure and experience variables	YES	YES
R-squared	0.33	0.32
Underidentification test (Kleibergen-Paap rk LM statistic):	150.02	129.847
Weak identification test (Cragg-Donald Wald F statistic):	175.24	153.18

Note: See notes of Table C.1.

The second stage regression results using this alternative instrument are presented in Table C.3. Consistent with our results in the main body of the text a one-year delay in the fertility time significantly reduces the amount of mismatch by about 6 percent of a standard deviation.

Table C.3. Second Stage Results for the Alternative Instrumental Variable: Miscarriage at the First Pregnancy

Measure of age at first birth	IV-1	IV-2
Relative age at the first birth	-0.0632 [0.0157]**	-0.0639 [0.0151]**
Absolute age at the first birth	-0.0665 [0.0164]**	-0.0674 [0.0159]**

Notes: The dependent variable is the rescaled total amount of mismatch. See notes to Table C.1 and Table 9.

[21] Following Miller (2011), we used questions on pregnancy losses from multiple interview rounds of the NLSY79 – the first round asked about the first pregnancy and subsequent rounds about pregnancies since the last interview – to fill in if there was a miscarriage at the first non-aborted pregnancy for women whose first pregnancy ended in abortion.

Appendix D: The Numbers behind the Figures

Flexibility and Mismatch

Table D.1. Mismatch, Fertility Timeline, and Occupation Flexibility (NLSY79)

	Male	Female
Flexibility score	-0.0365	-0.0236
	[0.0233]	[0.0244]
More than 3 years before the first birth	0.0444	0.030
	[0.0322]	[0.0425]
0-6 years after the first birth	-0.0411	0.0918
	[0.0258]	[0.0389]*
More than 6 years after the first birth	0.0056	0.178
	[0.0370]	[0.0490]**
More than 3 years before the first birth* Flexibility score	0.0355	-0.0104
	[0.0322]	[0.0382]
0-6 years after the first birth*Flexibility score	0.0415	0.0282
	[0.0283]	[0.0351]
More than 6 years after the first birth*Flexibility score	0.0456	0.0766
	[0.0320]	[0.0336]*
Observations	23261	18761

Notes: See notes to Table 10.

Table D.1 is the basis of our Figure 2. The last column of the table indicates that working in flexible occupations leads to a greater amount of mismatch for females after the first birth. Notice that in the last column, the coefficient for *More than 3 years before the first birth* Flexibility score* is negative but statistically insignificant, suggesting that long before females' first-birth working in flexible occupations does not yield greater mismatch. However, the effect becomes positive after the birth of the first child and in the case of *More than 6 years after the first birth* the flexibility interaction coefficient is positive and statistically significant. This means that mothers working in flexible occupations have much worse match quality compared with their childless counterparts more than 6 years following the birth of their first child – a one standard deviation increase in flexibility increases mismatch by about 8 percent of a standard deviation. This implies that mothers are trading off match quality for enhanced flexibility at work, an effect that becomes stronger as their children grow up. No such effects are observed for males.

Mismatch and the Gender Pay Gap

Table D.2 is the basis of our Figure 3. This table outlines the gender gap in the wage loss from mismatch (relative to the mean wage) for individuals with different levels of early match quality. We selected individuals in the sample based on their early match quality (i.e. an individual's match quality over the first 5 years of experience), distinguishing between the best and the worst occupational matches, respectively the top 10% and the bottom 10% in terms of match quality among both college graduates and non-graduates.

To illustrate the results given in this table, we provide the basis of the calculation for college graduates with 10 years of labor market experience. To simplify matters, we shall assume that total experience equals occupational tenure for everyone. Further, only the precisely estimated mismatch coefficients from the fourth column of Table 11 will be used. As there are no statistically significant gender differences in mismatch wage penalties, we base our calculation of mismatch and the gender wage gap on the following wage penalty coefficients: Current mismatch = -0.0311 (viz. *Mismatch*College and above*); Cumulative Mismatch= -0.0934 (viz. *Cumulative mismatch + Cumulative mismatch*College and above).* Current and past mismatch values are averages for these groups from data at the point of calculation. For example, for college graduates with the best early match quality, the average current mismatch is 1.29 for males and 2.14 for females at 10 years of experience. The average cumulative mismatch is 1.53 for males and 1.87 for females at 10 years of experience. Based on these values, we can compute the wage loss (relative to the mean wage) from each mismatch component. Thus, in the case of males, the wage effect of current mismatch is -0.04 (= -0.0311*1.29) and for cumulative mismatch it is -0.14 (= -0.0934*1.53). The corresponding wage losses for females are -0.07 (= -0.0311*2.14) and -0.17 (= -0.0934*1.87), respectively. The gender gap in wage loss associated with current mismatch – namely, the wage loss for females less the wage loss for males – is thus

-0.032 (the ninth column of Table D.2) and that associated with cumulative mismatch is also -0.026 (tenth column). The total gender gap in wage loss associated with mismatch is -0.058.

The penultimate column of Table D.2 presents a calculation of the gender gap in the wage loss from mismatch, using both the statistically insignificant and significant coefficients. We now base our calculations on the following wage penalty coefficients. Beginning with males: Current mismatch = -0.0262 (= 0.0049 + -0.0311) (viz. *Mismatch* + *Mismatch*College and above*); Cumulative Mismatch = -0.0934 (= -0.0421 + -0.0513) (viz. *Cumulative mismatch* + *Cumulative mismatch*College and above);* Current mismatch and occupational tenure interaction = 0.0005 (= -0.0009 + 0.0014) (viz. *Mismatch*Occupation tenure* + *Mismatch*Occupation tenure*College and above*). For females: Current mismatch = -0.0371 (= 0.0049 + 0.0063 + -0.0311 + -0.0172) (viz. *Mismatch* + *Mismatch*Female* + *Mismatch*College and above* + *Mismatch* College and above*Female*); Cumulative Mismatch = -0.0871 (= -0.0421 + -0.0052 + - 0.0513 + 0.0115) (viz. *Cumulative mismatch* + *Cumulative mismatch*Female* + *Cumulative mismatch*College and above* + *Cumulative mismatch*College and above*Female);* Current mismatch and occupational tenure interaction = -0.0016 (= -0.0009 + -0.0013 + 0.0014 + -0.0008) (viz. *Mismatch*Occupation tenure* + *Mismatch*Occupation tenure*Female* + *Mismatch*Occupation tenure*College and above* + *Mismatch*Occupation tenure*College and above* Female*). On this basis, the gender wage gap from current mismatch is equal to -0.045 (= -0.0371*2.14 + 0.0262*1.29), from current mismatch and occupational tenure interaction it is equal to -0.0407 (= -0.0016*10*2.14 - 0.0005*10*1.29), and from cumulative mismatch it is -0.0199 (= -0.0871*1.87 + 0.0934*1.53). The total gender wage gap attributable to mismatch is therefore -0.106.

In similar fashion, we can calculate wage losses for college graduates with the worst early match quality. The total gender gap in wage loss associated with mismatch is 0.006 using only the precisely estimated coefficients, and -0.037 using both significant and insignificant ones.

Table D.2. Mismatch and the Gender Wage Gap, by Experience, Early Match Quality, and Education

		Male				Female				Gender Gap			
				Wage Effect				Wage Effect					
		Cumulative mismatch	Current mismatch	Cumulative mismatch	Current mismatch	Cumulative mismatch	Current mismatch	Cumulative mismatch	Current mismatch	Cumulative mismatch	Current mismatch	Total (significant and insignificant coefficients)	Total (significant coefficients only)
College Graduates													
Experience and Early match quality	10 Year Exp.												
	Best	1.53	1.29	-0.14	-0.04	1.87	2.14	-0.17	-0.07	-0.032	-0.026	-0.106	-0.058
	Worst	3.71	2.05	-0.35	-0.06	3.65	2.05	-0.34	-0.06	0.006	0.000	-0.037	0.006
	20 Year Exp.												
	Best	1.8	1.31	-0.17	-0.04	2.36	2.01	-0.22	-0.06	-0.052	-0.022	-0.155	-0.074
	Worst	3.41	1.94	-0.32	-0.06	3.57	1.83	-0.33	-0.06	-0.015	0.003	-0.087	-0.012
	30 Year Exp.												
	Best	1.92	1.12	-0.18	-0.03	2.56	1.56	-0.24	-0.05	-0.060	-0.014	-0.164	-0.073
	Worst	3.37	1.84	-0.31	-0.06	3.57	2.11	-0.33	-0.07	-0.019	-0.008	-0.155	-0.027
Non-College Graduates													
Experience and Early match quality	10 Year Exp.												
	Best	1.56	1.43	-0.07	0.01	1.6	1.59	-0.07	0.01	-0.002	0.000	-0.021	-0.002
	Worst	3.96	2.53	-0.17	0.01	3.98	2.91	-0.17	0.01	-0.001	0.000	-0.043	-0.001
	20 Year Exp.												
	Best	2.01	1.68	-0.08	0.01	2.07	1.76	-0.09	0.01	-0.003	0.000	-0.049	-0.003
	Worst	4.08	2.38	-0.17	0.01	4.13	2.54	-0.17	0.01	-0.002	0.000	-0.076	-0.002
	30 Year Exp.												
	Best	2.17	1.43	-0.09	0.01	2.23	1.56	-0.09	0.01	-0.003	0.000	-0.068	-0.003
	Worst	4.11	2.32	-0.17	0.01	4.17	2.52	-0.18	0.01	-0.003	0.000	-0.111	-0.003

Notes: Early match quality is the determined according to an individual's match quality over first 5 years of experience. In this table we have the top decile (Best) and the bottom decile (Worst) early career matches. Current and past mismatch values are averages for these groups from data at the point of calculation. In the final column of the table we consider wage effects of mismatch using only the precisely estimated mismatch coefficients; for example, excluding the coefficient of the interaction between mismatch and occupation tenure. The penultimate column computes wages effects using both precisely and imprecisely estimated coefficient estimates.

Table D.3. Motherhood Wage Penalty

	Current Mismatch			Cumulative Mismatch			Mismatch Wage Effect (Significant only)			Mismatch Wage Effect (Significant and insignificant)			Wage Gap (Significant coefficients only)		Wage Gap (Significant and insignificant)	
Age	Birth Age=25	Birth Age=30	No Child	Birth Age=25	Birth Age=30	No Child	Birth Age=25	Birth Age=30	No Child	Birth Age=25	Birth Age=30	No Child	Birth Age=25	Birth Age=30	Birth Age=25	Birth Age=30
College Graduates																
24	**2.46**	1.72	2.52	**2.59**	2.54	2.65	**-0.23**	-0.20	-0.24	**-0.28**	-0.25	-0.29	**0.01**	0.04	**0.01**	0.04
25	**2.16**	1.70	2.01	**2.64**	2.58	2.68	**-0.21**	-0.19	-0.21	**-0.27**	-0.25	-0.27	**0.00**	0.02	**0.00**	0.02
26	**2.84**	1.97	1.88	**3.10**	2.59	2.65	**-0.25**	-0.20	-0.20	**-0.32**	-0.26	-0.26	**-0.06**	0.00	**-0.07**	0.00
27	**2.83**	1.97	2.01	**3.13**	2.60	2.70	**-0.25**	-0.19	-0.20	**-0.32**	-0.25	-0.26	**-0.05**	0.01	**-0.06**	0.01
28	**2.99**	1.98	2.08	**3.19**	2.63	2.77	**-0.25**	-0.19	-0.20	**-0.32**	-0.25	-0.26	**-0.05**	0.01	**-0.06**	0.01
29	**2.39**	**1.95**	2.06	**3.23**	**2.69**	2.81	**-0.23**	**-0.19**	-0.20	**-0.30**	**-0.25**	-0.26	**-0.03**	**0.01**	**-0.04**	**0.01**
30	**2.58**	**2.09**	1.83	**3.27**	**2.72**	2.78	**-0.22**	**-0.19**	-0.18	**-0.31**	**-0.25**	-0.25	**-0.04**	**0.00**	**-0.06**	**0.00**
31	2.72	**2.00**	1.75	3.40	**2.60**	2.72	-0.23	**-0.17**	-0.17	-0.32	**-0.24**	-0.24	-0.05	**0.00**	-0.07	**0.00**
32	2.58	**2.00**	1.63	3.43	**2.95**	2.81	-0.22	**-0.18**	-0.17	-0.31	**-0.26**	-0.24	-0.05	**-0.01**	-0.07	**-0.02**
33	2.40	**2.32**	1.73	3.47	**2.81**	2.81	-0.21	**-0.18**	-0.17	-0.30	**-0.26**	-0.24	-0.04	**-0.01**	-0.06	**-0.01**
34	2.42	**2.22**	1.97	3.52	**2.81**	2.94	-0.21	**-0.17**	-0.17	-0.30	**-0.25**	-0.25	-0.03	**0.00**	-0.05	**0.00**
35	2.61	**2.22**	1.91	3.60	**3.03**	3.01	-0.20	**-0.17**	-0.17	-0.31	**-0.26**	-0.25	-0.04	**0.00**	-0.06	**-0.01**
Non-College Graduates																
24	**1.99**	1.92	2.10	**2.06**	2.40	2.24	**-0.10**	-0.12	-0.11	**-0.10**	-0.12	-0.11	**0.01**	-0.01	**0.01**	-0.01
25	**2.01**	1.68	2.05	**2.11**	2.44	2.37	**-0.11**	-0.12	-0.12	**-0.10**	-0.12	-0.12	**0.01**	0.00	**0.01**	0.00
26	**2.39**	1.87	2.00	**2.52**	2.50	2.42	**-0.13**	-0.13	-0.12	**-0.13**	-0.13	-0.12	**-0.01**	0.00	**-0.01**	0.00
27	**2.53**	1.75	2.00	**2.71**	2.51	2.49	**-0.14**	-0.13	-0.13	**-0.14**	-0.13	-0.13	**-0.01**	0.00	**-0.01**	0.00
28	**2.14**	1.81	2.02	**2.71**	2.65	2.58	**-0.14**	-0.13	-0.13	**-0.14**	-0.14	-0.14	**-0.01**	0.00	**-0.01**	0.00
29	**2.14**	**1.84**	2.02	**2.72**	**2.69**	2.61	**-0.14**	**-0.14**	-0.13	**-0.15**	**-0.14**	-0.14	**-0.01**	**0.00**	**-0.01**	**0.00**
30	**2.24**	**1.91**	1.92	**2.83**	**2.72**	2.63	**-0.14**	**-0.14**	-0.13	**-0.15**	**-0.15**	-0.14	**-0.01**	**0.00**	**-0.01**	**0.00**
31	2.14	**1.88**	1.97	2.87	**2.73**	2.66	-0.15	**-0.14**	-0.14	-0.16	**-0.15**	-0.15	-0.01	**0.00**	-0.01	**0.00**
32	2.43	**2.00**	1.86	3.02	**2.85**	2.67	-0.15	**-0.15**	-0.14	-0.17	**-0.16**	-0.15	-0.02	**-0.01**	-0.02	**-0.01**
33	2.32	**2.23**	1.82	3.06	**2.78**	2.67	-0.16	**-0.14**	-0.14	-0.17	**-0.16**	-0.15	-0.02	**-0.01**	-0.02	**-0.01**
34	2.35	**2.00**	1.90	3.32	**2.69**	2.71	-0.17	**-0.14**	-0.14	-0.19	**-0.16**	-0.16	-0.03	**0.00**	-0.04	**0.00**
35	2.63	**2.46**	1.93	3.34	**2.93**	2.89	-0.17	**-0.15**	-0.15	-0.20	**-0.18**	-0.17	-0.02	**0.00**	-0.03	**-0.01**

Mismatch, Age at First Birth, and the Motherhood Wage Penalty

Table D.3 is the basis of Figure 4. This table provides the differences in wage loss associated with mismatch for females who gave birth to their first child either at age 25 or age 30 vis-à-vis their childless counterparts.

To illustrate the results given in this table, we provide the basis of the calculation for college graduates entering the labor market at age 23. Again, we first consider the effects for precisely estimated mismatch coefficients shown in the final column of Table 11: *Cumulative mismatch* (-0.0509), *Mismatch*College and above* (-0.0429), and *Mismatch*College and above*Occupation tenure* (0.0029). Current and cumulative mismatch values are averages for these groups from data at the point of calculation. For example, at the age of 30 (or 5 years after the first birth for those who had a child at 25), the current mismatch is 1.83 for those without a child, 2.58 for those whose first birth age is 25, and 2.09 for those whose first birth age is 30; the cumulative mismatch is 2.78 for those without a child, 3.27 for mothers whose first birth age is 25, and 2.72 for mothers whose first birth age is 30. Based on these values, we can compute the wage loss (relative to the mean wage) due to each mismatch component.

The wage effect of current mismatch is -0.0785 (= -0.0429*1.83) for those without a child, -0.111 (=-0.0429*2.58) for mothers whose first birth age is 25, and -0.090 (= -0.0429*2.09) for mothers whose first birth age is 30. The wage effect of cumulative mismatch is -0.142 (= -0.0509*2.78) for those without a child, -0.166 (=-0.0509*3.27) for mothers whose first birth age is 25, and -0.138 (= -0.0509*2.72) for mothers whose first birth age is 30. With respect to occupational tenure, by the age of 30 all groups will have 7 years of occupational tenure (here, for purposes of illustration we are imposing the restriction that they are employed in the same occupation). For its part, the wage effect from the interaction of current mismatch and occupational tenure among college graduates is 0.037 (= 0.0029*1.83*7) for those without a child, 0.052 (=0.0029*2.58*7) for mothers whose first birth age is 25, and 0.042 (= 0.0029*2.09*7) for mothers whose first birth age is 30. At the age of 30, the differences in wage loss associated with mismatch between mothers whose first birth age is 25 and their childless counterparts is -0.042 [=(-0.0429*2.58+-0.0509*3.27+ 0.0029*2.58*7) - (-0.0429*1.83+-0.0509*2.78+ 0.0029*1.83*7)]. Similarly, the differences in wage loss associated with mismatch, between mothers to be whose first birth age is 30 and their childless counterparts at the at the same age is 0.002 [= (-0.0429*2.09 + -0.0509*2.72 + 0.0029*2.09*7) - (-0.0429*1.83+-0.0509*2.78+ 0.0029*1.83*7)].

Figure 4 graphs these losses over the timeline of first birth (comparing the highlighted portions of the penultimate set of two columns). Comparing the two groups of mothers (with first birth age at 25 and 30) 5 years after the first birth will require us to compare their wage losses at ages 30 and 35, respectively. The gap for mothers who gave first birth at age 25, relative to their childless counterparts was earlier calculated to be -0.042. This gap is 0 for mothers who gave first birth at age 30 relative to their childless counterparts at age 35. This implies more than 4 percentage points higher wage loss due to mismatch for individuals with earlier births, 5 years after birth. Figure 4 captures only the gaps calculated using the precisely estimated coefficients. The last two columns of Table D.3 provide the corresponding gaps calculated using all relevant coefficients (significant and non-significant).

Appendix E: Marginal Effects for Table 12

Table E1. Pairwise Comparisons of Predictive Margins

Comparison groups	Difference in Predicted Probabilities	Std Err.	95% Confidence Interval	
Non-college graduate NLSY79 female vs non-college graduate NLSY79 male	-0.0138	0.0035	-0.0205	-0.007
College graduate NLSY79 male vs non-college graduate NLSY79 male	-0.0183	0.0050	-0.0280	-0.0086
College graduate NLSY79 female vs non-college graduate NLSY79 male	-0.0244	0.0054	-0.0351	-0.0137
Non-college graduate NLSY97 male vs non-college graduate NLSY79 male	0.0139	0.0044	0.0052	0.0226
Non-college graduate NLSY97 female vs non-college graduate NLSY79 male	-0.0107	0.0047	-0.0200	-0.0015
College graduate NLSY97 male vs non-college graduate NLSY79 male	0.0171	0.0064	0.0046	0.0297
College graduate NLSY97 female vs non-college graduate NLSY79 male	0.0102	0.0061	-0.0018	0.0222
College graduate NLSY79 male vs non-college graduate NLSY79 female	-0.0045	0.0054	-0.0151	0.0061
College graduate NLSY79 female vs non-college graduate NLSY79 female	-0.0106	0.0058	-0.0221	0.0008
Non-college graduate NLSY97 male vs non-college graduate NLSY79 female	0.0277	0.0048	0.0183	0.037
Non-college graduate NLSY97 female vs non-college graduate NLSY79 female	0.0030	0.0050	-0.0068	0.0128
College graduate NLSY97 male vs non-college graduate NLSY79 female	0.0309	0.0068	0.0176	0.0442
College graduate NLSY97 female vs non-college graduate NLSY79 female	0.0239	0.0065	0.0112	0.0367
College graduate NLSY79 female vs college graduate NLSY79 male	-0.0061	0.0059	-0.0176	0.0055
Non-college graduate NLSY97 male vs college graduate NLSY79 male	0.0322	0.0059	0.0206	0.0439
Non-college graduate NLSY97 female vs college graduate NLSY79 male	0.0076	0.0062	-0.0046	0.0197
College graduate NLSY97 male vs college graduate NLSY79 male	0.0354	0.0070	0.0218	0.0491
College graduate NLSY97 female vs college graduate NLSY79 male	0.0285	0.0067	0.0154	0.0416
Non-college graduate NLSY97 male vs college graduate NLSY79 female	0.0383	0.0060	0.0266	0.05
Non-college graduate NLSY97 female vs college graduate NLSY79 female	0.0136	0.0063	0.0013	0.0259
College graduate NLSY97 male vs college graduate NLSY79 female	0.0415	0.0071	0.0276	0.0554
College graduate NLSY97 female vs college graduate NLSY79 female	0.0346	0.0068	0.0213	0.0478
Non-college graduate NLSY97 female vs non-college graduate NLSY97 male	-0.0247	0.0042	-0.0330	-0.0163
College graduate NLSY97 male vs non-college graduate NLSY97 male	0.0032	0.0060	-0.0086	0.015
College graduate NLSY97 female vs non-college graduate NLSY97 male	-0.0037	0.0057	-0.0148	0.0073
College graduate NLSY97 male vs non-college graduate NLSY97 female	0.0279	0.0063	0.0155	0.0403
College graduate NLSY97 female vs non-college graduate NLSY97 female	0.0209	0.0060	0.0091	0.0327
College graduate NLSY97 female vs college graduate NLSY97 male	-0.0070	0.0066	-0.0200	0.006

Notes: These differences in predicted probabilities are generated using Table 12 IV-Probit results. Stata 15 *margins* command is used with *pwcompare* option calculated at specified values of education category, cohort and gender. Significant differences are highlighted.

TABLES

Table 1. Sample Construction

	NLSY79 (1979-2014)				NLSY97 (1997-2013)			
	Remaining		Remaining		Remaining		Remaining	
Criterion for sample selection	Male	Female	Male	Female	Male	Female	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 in 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 years	2,146	2,217	26,645	21,144	2,288	2,045	13,356	10,098
9 Have valid ability measures	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

Note: The values refer to annual data and are not derived from 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
Highest education = high school	0/1 Dummy	0.45	0.46	0.45	0.33	0.37	0.28
Highest education > high school	0/1 Dummy	0.19	0.17	0.22	0.21	0.21	0.22
Highest education ≥ 4-year college	0/1 Dummy	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 (married, divorced, widowed, seperated)	0/1 Dummy	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			
Total labor market experience (mean)	Mean years worked	14.3	14.4	14.2	8.75	8.67	8.85
Total labor market experience (median)	Median years worked	13.0	13.0	12.0	9.00	9.00	9.00
Occupational tenure (mean)	Mean years worked in the same occupation	6.36	6.35	6.37	3.52	3.50	3.54
Occupational tenure (median)	Median years worked in the same occupation	4.00	4.00	4.00	3.00	3.00	3.00

Table 3. Descriptive Statistics of Worker Skill Endowments

		NLSY79					NLSY97				
		# Individuals	Math	Verbal	STM	Social	# 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 and above	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 for each skill type are reported. Educational groups are defined by the highest degree ever completed. Standard errors are given in brackets.

Table 4. Extent and Size of Mismatch, by Experience, Skill Type, and Gender

			[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 the measure of the gap between the workers' endowments and occupational requirements of their jobs when the the average skill endowment exceeds the average occupational skill requirements. "Endowment > requirement" is a crude measure of over-qualification and the table reports the percentage of workers with endowments that are greater than the skill levels required by their occupation. "Share of over-qualified" refers to the share of workers who are more than one standard deviation more endowed than required. Measures of under-qualification are similarly constructed.

Table 5. The Determinants of Mismatch: The Role of Gender and Education (NLSY79)

	OLS				IV		
	(1)	(2)	(3)	(4)	(5)	Full-Timers	Part-Timers
Female	-0.0253	-0.0712	-0.0309	-0.0325	-0.0310	-0.0402	-0.0017
	[0.0256]	[0.0299]*	[0.0278]	[0.0277]	[0.0279]	[0.0298]	[0.0389]
Female* College Graduate		0.1588	0.1536	0.1509	0.1557	0.1515	-0.0709
		[0.0569]**	[0.0509]**	[0.0508]**	[0.0510]**	[0.0532]**	[0.1003]
College Graduate		-0.1684	-0.2624	-0.2649	-0.2642	-0.2429	-0.1907
		[0.0384]**	[0.0388]**	[0.0389]**	[0.0388]**	[0.0393]**	[0.0889]*
Additional demographic and human capital controls and occupational requirements	NO	NO	YES	YES	YES	YES	YES
Tenure and experience variables	NO	NO	NO	YES	YES	YES	YES
Observations	42022	42022	42022	42022	42022	35893	6129

Notes: The dependent variable is the rescaled total amount of mismatch. The additional variables in column (3) are race and average measures of individual skills and occupational requirements. In addition, column (4) includes employer tenure, occupational tenure and total experience, and their squared values, and also the interaction terms between skills and occupational tenure variables and between occupational requirements and occupational tenure variables. Our baseline models will include all the controls in column 4. In column (5) tenure and experience variables, and their interaction terms are instrumented. Clustered standard errors are given in brackets. **, * denote significance at the 0.01 and 0.05 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]*	[0.0239]	[0.0243]	[0.0345]**	[0.0247]	[0.0252]	[0.0355]**
Female * Have at least one child	0.0971			0.1305			0.1393		
	[0.0433]*			[0.0392]**			[0.0403]**		
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]	[0.0258]	[0.0264]	[0.0339]	[0.0264]	[0.0269]	[0.0349]
Female * Ever married	0.034			-0.0576			-0.0492		
	[0.0477]			[0.0398]			[0.0408]		
Observations	42022	23261	18761	42022	23261	18761	42022	23261	18761

Notes: The dependent variable is the rescaled total amount of mismatch. All specifications use the full set of controls as described in column (4) of Table 5. In the IV specifications, the tenure and experience variables and their interaction terms are instrumented. The standard errors in brackets are clustered at the individual level. **, * denote significance at the 0.01 and 0.05 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 an individual's average total amount of skill surplus exceeds the average of occupational requirements by 1 standard deviation. Conditional fixed-effects logistic regression estimates are reported. All specifications contain the full set of controls. Standard errors are given in brackets. + denotes significance at the 0.1 level.

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]	[0.0313]	[0.0387]	[0.0319]	[0.0393]
0-6 years after the first birth	-0.0429	0.0567	-0.0375	0.0962	-0.0442	0.0829
	[0.0324]	[0.0393]	[0.0252]	[0.0339]**	[0.0257]+	[0.0349]*
More than 6 years after the first birth	-0.003	0.1043	0.015	0.1731	0.0022	0.1512
	[0.0391]	[0.0405]*	[0.0358]	[0.0459]**	[0.0367]	[0.0470]**
Observations	23261	18761	23261	18761	23261	18761

Notes: The dependent variable is the rescaled total amount of mismatch. The baseline group comprises individuals who are within the 3 years period before the first birth. All specifications use the full set of controls. In the IV specification, tenure and experience variables and their interaction terms are instrumented. Standard errors given in brackets are clustered at the individual level. **, *, + denote significance at 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	Female	Male	Female
Relative age at the first birth	0.0038	-0.004	-0.0398	0.0038	-0.0361
	[0.0013]**	[0.0015]**	[0.0166]*	[0.0032]	[0.0160]*
Absolute age at the first birth	0.0031	-0.0038	-0.0366	0.0028	-0.0332
	[0.0013]*	[0.0017]*	[0.0151]*	[0.0033]	[0.0146]*

Notes: The dependent variable is the rescaled total amount of mismatch. Relative birth age is calculated as "the year of the first birth - the year of labor market entry". All specifications include the full set of controls. In IV-1, individuals' relative/absolute birth ages are instrumented using their siblings' average age at the first birth. Only female estimates are reported as endogeneity of birth age is not an issue in the case of males. In IV-2, tenure and experience variables are also instrumented. Robust standard errors are given in brackets. **, * denote significance at the 0.01 and 0.05 levels, respectively.

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

	Male	Female
Have at least one child	-0.0199	0.1185
	[0.0253]	[0.0373]**
Have at least one child* Flexibility score	0.0409	0.0636
	[0.0242]+	[0.0280]*
Flexibility score	-0.0247	-0.0238
	[0.0169]	[0.0210]

Notes: The dependent variable is the rescaled total amount of mismatch. All models include full set of controls. Reported coefficients are from fixed effects models where tenure and experience variables are instrumented. Models without instrumentation proudce almost identical coefficient estimates. Standard errors given in brackets are clustered at the individual level. **, *, + denote significance at the 0.01, 0.05 and 0.1 levels, respectively.

Table 11. Mismatch and Wage Outcomes (NLSY79)

	IV					
	All				Male	Female
Mismatch	-0.0236	-0.0213	-0.0038	0.0049	0.0062	0.0038
	[0.0062]**	[0.0082]**	[0.0080]	[0.0090]	[0.0105]	[0.0095]
Mismatch*Occupation tenure		-0.0006	-0.0005	-0.0009	-0.0006	-0.0012
		[0.0016]	[0.0016]	[0.0017]	[0.0023]	[0.0021]
Cumulative mismatch			-0.0598	-0.0421	-0.0384	-0.0509
			[0.0078]**	[0.0094]**	[0.0103]**	[0.0081]**
Female	-0.1429	-0.1416	-0.1449	-0.1257		
	[0.0196]**	[0.0196]**	[0.0269]**	[0.0267]**		
Female*Mismatch	0.0014	0.0056	0.0019	0.0063		
	[0.0080]	[0.0089]	[0.0088]	[0.0102]		
Female*Mismatch*Occupation tenure		-0.0014	-0.0015	-0.0013		
		[0.0012]	[0.0012]	[0.0015]		
Female*Cumulative mismatch			0.005	-0.0052		
			[0.0093]	[0.0108]		
Mismatch*College and above				-0.0311	-0.0362	-0.0429
				[0.0143]*	[0.0152]*	[0.0134]**
Mismatch*Occupation tenure*College and above				0.0014	-0.0004	0.0029
				[0.0021]	[0.0023]	[0.0017]+
Cumulative mismatch*College and above				-0.0513	-0.0639	-0.0217
				[0.0130]**	[0.0156]**	[0.0150]
Female*Mismatch*College and above				-0.0172		
				[0.0181]		
Female*Mismatch*Occupation tenure*College and above				-0.0008		
				[0.0025]		
Female*Cumulative mismatch*College and above				0.0115		
				[0.0134]		
Observations	42022	42022	42022	42022	23261	18761

Notes: The dependent variable is the log real wage (measured in 2002 dollars). Estimates from IV models using the full set of controls are reported. Tenure and experience variables are instrumented. The standard errors in brackets are clustered at the individual level. **, *, + denote significance at the 0.01, 0.05 and 0.1 levels, respectively.

Table 12. The Determinants of Mismatch and Over-qualification: Cohort Differences (NLSY79 and NLSY97)

	IVREG	IVPROBIT
Female	-0.0386	-0.1554
	[0.0142]**	[0.0402]**
Female*NLSY97	-0.0738	-0.1226
	[0.0225]**	[0.0630]+
Female*College graduate	0.102	0.0654
	[0.0281]**	[0.0795]
Female*College graduate*NLSY97	0.0328	0.1238
	[0.0429]	[0.1196]
College graduate*NLSY97	0.181	0.2489
	[0.0304]**	[0.0836]**
College graduate	-0.2304	-0.2125
	[0.0204]**	[0.0577]**
NLSY97	0.1029	0.1952
	[0.0181]**	[0.0514]**
Observations	39112	39112

Notes: The dependent variables are the rescaled total amount of mismatch and a dummy indicator for over-qualification as defined in Table 7. NLSY97 is a dummy set equal to 1 if the observation is from the NSLY97 cohort. Tenure and experience variables and their interaction terms are instrumented. Robust standard errors are given in brackets.**, + denote significance at the 0.01 and 0.1 levels, respectively.

Table 13. The Determinants of Mismatch: The Role of Family, Fertility Timeline, and Cohort

	Male	Female	Male	Female
Have at least one child	-0.0261	0.0842		
	[0.0198]	[0.0284]**		
Have at least one child*NLSY97	0.0564	-0.0803		
	[0.0330]+	[0.0438]+		
More than 3 years before the first birth			0.0379	0.0353
			[0.0263]	[0.0307]
0-6 years after the first birth			-0.0195	0.0588
			[0.0230]	[0.0295]*
More than 6 years after the first birth			0.0667	0.1635
			[0.0358]+	[0.0442]**
More than 3 years before the first birth*NLSY97			0.0007	0.0472
			[0.0417]	[0.0482]
0-6 years after the first birth*NLSY97			0.0519	-0.0276
			[0.0379]	[0.0463]
More than 6 years after the first birth*NLSY97			-0.0339	-0.0825
			[0.0591]	[0.0685]
Observations	22152	16960	22152	16960

Notes: The dependent variable is the rescaled total amount of mismatch. All specifications use the full set of controls. Reported coefficients are from IV-Fixed Effects models where the tenure and experience variables are instrumented. Standard errors given in brackets. **, *, + denote significance at the 0.01, 0.05 and 0.1 levels, respectively.

Table 14. Mismatch, Fertility, and Breadwinners (NLSY79 and NLSY97)

NLSY79	IV		FE		IV-FE	
	Female	Male	Female	Male	Female	Male
Breadwinner*Have at least one child	0.0955	-0.0412	0.0512	-0.0738	0.0528	-0.0742
	[0.0500]+	[0.0649]	[0.0331]	[0.0390]+	[0.0416]	[0.0481]
Breadwinner	-0.083	-0.1008	-0.0763	0.0145	-0.0766	0.0104
	[0.0412]*	[0.0474]*	[0.0281]**	[0.0306]	[0.0347]*	[0.0350]
Have at least one child	0.0229	0.03	0.0824	0.0517	0.0591	0.0464
	[0.0467]	[0.0645]	[0.0315]**	[0.0391]	[0.0436]	[0.0519]
Observations	14039	15998	14039	15998	14039	15998

NLSY97						
Breadwinner*Have at least one child	-0.0195	0.1543	0.0014	0.0184	0.0039	0.0157
	[0.0935]	[0.1231]	[0.0614]	[0.0754]	[0.0763]	[0.0848]
Breadwinner	-0.0113	-0.074	-0.0878	-0.07	-0.0834	-0.0655
	[0.0668]	[0.0826]	[0.0453]+	[0.0566]	[0.0553]	[0.0685]
Have at least one child	0.0279	-0.2038	0.0319	-0.0002	0.0373	0.0071
	[0.0740]	[0.1185]+	[0.0564]	[0.0736]	[0.0846]	[0.0875]
Observations	2282	2384	2282	2384	2282	2384

Note: The dependent variable is the rescaled total amount of mismatch. All specifications use the full set of controls with instrumentation for tenure and experience variables and their interactions. Only married individuals who are living with a spouse are included in the sample. Breadwinner status is determined when an individual's spouse does not have any wage income or the spousal wage income is less than that of the respondent. Standard errors reported in brackets are clustered at the individual level. **, + denote significance at the 0.01 and 0.1 levels, respectively.

Table 15. The 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 and above	-0.2347	-0.2228	-0.1379	-0.12	-0.1075	-0.1142	-0.1102	-0.1065	-0.0468	0.003	0.0126	-0.0652
	[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 child*College and above	0.1063			0.0316			0.0621			-0.0494		
	[0.0492]*			[0.0756]			[0.0768]			[0.0837]		
College and above	-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 and above	-0.5384	-0.8614	0.5685	-0.3057	-0.3034	0.1596	-0.0842	-0.1537	0.0999	-0.2732	-0.0886	-1.0051
	[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 child*College and above	0.9308			0.4983			0.0592			-0.3224		
	[0.4244]*			[0.4383]			[0.3183]			[0.3821]		
College and above	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: The dependent variables are the rescaled total amount of mismatch and a dummy indicator for over-qualification by each skill; for definition of overqualification see Table 7. All specifications use the full set of controls. reported coeficients are from IV-Fixed Effects models where the tenure and experience variables are instrumented. Standard errors are given in brackets. **, *, + denote significance at the 0.01, 0.05 and 0.1 levels, respectively.

Table 16. Mismatch and Wages by Skill Type (NLSY79)

	IV						IV-FE	
	Male	Female	Male	Female	Male	Female	Male	Female
Math mismatch	-0.0141	-0.0073	-0.0122	0.0009	-0.0153	-0.0076	-0.009	-0.0091
	[0.0083]+	[0.0075]	[0.0110]	[0.0100]	[0.0118]	[0.0101]	[0.0079]	[0.0060]
Verbal mismatch	-0.0078	-0.0048	0.0027	0.0031	0.0152	0.0136	-0.0015	0.0061
	[0.0077]	[0.0075]	[0.0104]	[0.0095]	[0.0117]	[0.0098]	[0.0081]	[0.0069]
STM mismatch	-0.0028	-0.0055	-0.0098	-0.0243	-0.0043	-0.0207	-0.0087	-0.0063
	[0.0070]	[0.0076]	[0.0086]	[0.0097]*	[0.0085]	[0.0090]*	[0.0063]	[0.0065]
Social-noncognitive mismatch	0.0054	0.0037	0.0013	0.0006	-0.0027	-0.0049	-0.0053	-0.0009
	[0.0065]	[0.0060]	[0.0082]	[0.0076]	[0.0091]	[0.0078]	[0.0072]	[0.0059]
Math mismatch*Occupation tenure			-0.0007	-0.0025	-0.0006	-0.0024	-0.001	-0.0015
			[0.0026]	[0.0018]	[0.0026]	[0.0018]	[0.0019]	[0.0011]
Verbal mismatch*Occupation tenure			-0.0036	-0.0026	-0.0034	-0.0026	-0.0037	-0.0032
			[0.0027]	[0.0020]	[0.0026]	[0.0020]	[0.0020]+	[0.0014]*
STM mismatch*Occupation tenure			0.0023	0.0057	0.0024	0.0058	0.0018	0.0018
			[0.0017]	[0.0018]**	[0.0017]	[0.0018]**	[0.0013]	[0.0014]
Social-noncognitive mismatch*Occupation tenure			0.0012	0.0007	0.0011	0.0008	0.0003	-0.0001
			[0.0019]	[0.0014]	[0.0019]	[0.0014]	[0.0016]	[0.0011]
Cumulative math mismatch					0.0041	0.0185	-0.0464	-0.0224
					[0.0123]	[0.0109]+	[0.0139]**	[0.0132]+
Cumulative verbal mismatch					-0.0241	-0.0239	-0.0438	-0.0228
					[0.0118]*	[0.0102]*	[0.0152]**	[0.0129]+
Cumulative STM mismatch					-0.0135	-0.0081	0.0081	0.000
					[0.0084]	[0.0095]	[0.0094]	[0.0113]
Cumulative social-noncognitive mismatch					0.0085	0.0114	-0.0095	-0.0215
					[0.0089]	[0.0080]	[0.0104]	[0.0102]*
Observations	23261	18761	23261	18761	23261	18761	23261	18761

Notes: The dependent variable is the log real wage (measured in 2002 dollars). Only coefficient estimates for the mismatch variables are reported. All specifications use the full set of controls. Reported coefficients are from IV and IV-Fixed Effects models where the tenure and experience variables are instrumented. Robust standard errors are given in brackets. **, *, + denote significance at the 0.01, 0.05 and 0.1 levels, respectively.