Educational Mismatch and the Earnings Distribution: Where Does the Mismatch Bite?

Keith A. Bender[1]
University of Aberdeen

Kristen Roche[2]
Mount Mary University

Abstract

This paper focuses on the interrelationship between educational mismatch and earnings taking two new approaches. First, we examine decompositions of the mismatch wage gap, finding that characteristics explain less than half of the mismatch penalty. Second, we use quantile regression to examine the mismatch penalty across the earnings distribution, showing that the penalty shrinks as the position in the earnings distribution increases. Different reasons for mismatch show heterogeneity in the penalty at the mean and at different points in the distribution with larger penalties for being mismatched due to working conditions, location, family, and no available job.

JEL: I2, J16, J24, J31

Keywords: educational mismatch; earnings; decomposition; quantile regression

Draft: January 2017

[1] Department of Economics, University of Aberdeen, Aberdeen AB24 3QY, Scotland, UK; kabender@abdnac.uk; 44-(0)1224-273411.

[2] Business Administration Department, Mount Mary University, 2900 North Menomonee River Pkwy, Milwaukee, WI, 53222; rochek@mtmary.edu; (414) 930-3419.

INTRODUCTION

College students invest in their education with the assumption that the knowledge and skills they acquire in their degree will be useful and demanded in the labor market. Indeed, in an efficient labor market, the human capital that workers possess would match the human capital that firms require. Yet research finds that there is a mismatch between the human capital workers acquire through education and the human capital required for the job. Estimates vary, but generally about 15 to 30% of workers are educationally mismatched in developed economies (e.g. Bender and Roche, 2013; Chevalier, 2003; Wolbers, 2003).

While there is some debate over the causes of mismatch and whether it is a labor supply or labor demand phenomenon (or both), there is consistent and robust empirical evidence that mismatch is correlated with adverse labor market outcomes. These include lower job satisfaction (e.g. Baker *et al*. 2010; Bender and Heywood, 2006), more turnover (e.g. Bender and Heywood, 2009; Wolbers, 2003), and lower pay (e.g. Bender and Heywood, 2009; Chevalier 2003).

This paper investigates the latter issue by focusing on the interrelationship between educational mismatch and earnings. Using data on full-time workers with college degrees in the U.S., we estimate a 7% wage penalty for workers whose degrees are somewhat related to their job, and a 22% wage penalty for workers who degrees are not at all related to their job on average, *ceteris paribus*. For the latter group, wage penalties vary greatly depending on the reason for the worker's mismatch. For example, workers who are mismatched due to location, family responsibilities, and limited job availability incur large penalties exceeding 30%. However, if a worker is mismatched due to pay or a career change, the penalty is smaller.

These findings, discussed in more detail below, are fairly representative of the earnings penalties found in the literature. In this paper, we add to the current literature using two new approaches in investigating the mismatch-earnings relationship. First, we decompose the earnings differential to estimate the proportion of the penalty that is attributable to observed characteristics versus unobserved characteristics. Using an Oaxaca-style decomposition, we estimate that observed characteristics, such as gender, age, occupation, and industry, explain less than half of the differential. In addition, observed characteristics explain a greater portion of the differential for males relative to females.

Second, we examine the mismatch penalty across the earnings distribution. Using a mean squared error (MSE) decomposition which attempts to capture a more detailed measure of differences in earnings between the matched and mismatched by comparing the first two moments of the two distributions, our results suggest that earnings differences between the matched and mismatched are driven more by differences in the distributions than by differences in average earnings. We also find that it is the difference in distributions that explains a larger percentage of the differential for men compared to women. Next, we employ quantile regression to estimate the differential across the earnings distribution. Among workers with jobs that are not at all related to their education, the penalty decreases as the position in the earnings distribution increases. These penalties differ by gender and by reason for mismatch.

This paper is organized as follows. First, a brief review of the literature provides a background for the research on quantile regression and educational mismatch. The data are then defined and examined for descriptive statistics and results are presented on the earnings differential decomposition and the earnings differential across the earnings distribution. Finally, the results of the paper are summarized and recommendations for future work are made.

I. LITERATURE REVIEW

Educational mismatch

Previous research on educational mismatch focuses on the effects of being employed in a job that is not well matched with a worker's education. There is consistent and robust empirical evidence in the economics literature that educational mismatch is correlated with adverse labor market outcomes. These include lower job satisfaction (e.g. Baker *et al*. 2010; Bender and Heywood, 2006), more turnover (e.g. Bender and Heywood, 2009; Wolbers 2003), and lower pay (e.g. Bender and Heywood, 2009; Chevalier, 2003). Given that our paper investigates the latter outcome, we focus our literature review on the educational mismatch earnings penalty and its measurement.

The literature defines educational mismatch as vertical mismatch (over- or under-education) or horizontal mismatch (the match between worker skills or education and the job being done). In general, wage penalties are estimated with both types of mismatch, although the magnitudes of the penalties differ depending on the type. For example, a meta-analysis by Groot and Maasen van den Brink (2000) shows that the over-educated experience a 14% earnings penalty. Another study using UK data (Chevalier, 2003) estimates a 5-11% penalty for mismatched workers who have similar unobserved skills as matched workers, and a 22-26% penalty for mismatched workers with lower skill endowments as matched workers. Studies that investigate the horizontal mismatch penalty (see Borghans and de Grip, 2000; Bender and Heywood, 2009; Bender and Roche, 2013) find a similar result – working in a job unrelated to one's field of education is associated with significantly diminished earnings, and these penalties increase with the severity of mismatch.

However, these studies above examine the penalty using standard linear regression methods 0 essentially estimating the penalty at the conditional mean. More recently, economists have started to examine mismatch penalties conditional on the position in the earnings distribution. Thus, a brief review of quantile regression and its application to educational mismatch follows.

Quantile regression

Previous research that estimates wage equations finds that wages vary significantly across the distribution and therefore estimation of wage determinants using OLS can provide biased results. Seminal papers focus on measuring the returns to education. Buchinsky (1994 and 1998) implements quantile regression to estimate the returns to skills, i.e., the return to education and the return to experience, across the wage distribution. Buchinsky finds that relative to median regression, OLS underestimates the returns to skills, and the returns to skills increase across the wage distribution. That is, workers at the bottom of the wage distribution experience smaller returns to education and experience, and workers at the top of the wage distribution experience higher returns to education and experience.

In a novel paper using twins data, Arias *et al*. (2000) estimate the returns to education using instrumental variables quantile regression. The authors regard the wage distribution as a reflection of the range of unobservable ability, so that people at the bottom of the wage distribution are believed to have less ability and people at the top of the wage distribution are believed to have more ability. Accordingly, they argue that education and unobserved ability have a complementary relationship, and given the additional indirect effect of education on human capital, education helps high-ability individuals more.

Subsequent research expands on the usefulness and application of quantile regression. Yu *et al*. (2003) is a highly-cited paper that summarizes the motivation and many applications of quantile regression. Recently, quantile regression has been applied to research on educational mismatch. The results from these papers, which mainly measure the effect of mismatch on wages in developed economies, are mixed.

Evidence from Spain (Budria and Moro-Egido, 2008) and Northern Ireland (McGuinness and Bennett, 2007) show a mismatch penalty that narrows as the position in the earnings distribution increases. However, in an opposing paper, Hernández and Serrano (2012) find that the mismatch penalty in Spain is larger for high-wage workers in the upper part of the distribution, implying that it is not unobservable characteristics, but rather educational mismatch itself driving the wage inequality.[1]

In summary, further analysis is warranted to better understand the interrelationship between educational mismatch and earnings, particularly in a distributional context. While the current literature agrees that mismatch is correlated with lower pay, it is conflicted on how the earnings penalty differs at different positions in the earnings distribution. Furthermore, we know little of what explains the mismatch penalty, and whether it is attributable to a difference in distributions or a difference in means between matched and mismatched workers.

II. DATA

This paper uses the 2003 US National Survey of College Graduates (NSCG) dataset from the US National Science Foundation (NSF). The data are a nationally representative sample of approximately 66,000 university educated workers who either work in science, technology, engineering, or math (STEM) fields, or have earned at least a bachelor's degree in a STEM field.

The NSCG asks respondents a key subjective question giving a measure of horizontal mismatch, "Thinking about the relationship between your work and your education, to what extent is your work related to your highest degree? Closely related, somewhat related, or not at all related." We identify these workers as closely matched, moderately mismatched, and severely mismatched, respectively. If workers are severely mismatched, the NSCG asks them follow-up questions about the most important reason for working in a field that is not at all related to their highest degree. Reasons include pay and promotion opportunities, working conditions, job location, change in career or professional interests, family-related reasons, job not available, and other.[2] Table 1 presents rates of mismatch, rates for the reasons of severely mismatched workers, and mean and median earnings by mismatch type. It is interesting to note that there is an association between increasing mismatch and lower mean and median earnings for the overall sample as well as in both the female and male samples.

(Table 1 here)

In addition to the educational mismatch variables, the dataset provides a standard set of socioeconomic variables. The following analysis restricts the data to full-time workers who report positive earnings in order to examine workers in career-type jobs only.

III. Results

Mismatch penalty using Ordinary Least Squares (OLS)

First, we examine the mismatch penalty using the traditional approach, ordinary least squares (OLS). Log hourly earnings are regressed on indicators for moderate mismatch, severe mismatch, and a standard set of covariates, including gender, age, age squared, race, marital status, US citizenship, experience, experience squared, tenure, tenure squared, highest degree, occupation, industry, disability status, employment sector, firm size, and region.

Similar to previous research, we find fairly substantial earnings penalties using OLS as seen in Table 2. Relative to matched workers, *ceteris paribus*, moderate mismatch is associated with a 6.8% penalty and severe mismatch is associated with a 22.4% penalty. These penalties differ only slightly by gender. For severely mismatched workers, the mean earnings penalty varies depending on the most important reason for their mismatch, as shown in the separate regression results in the bottom panel of Table 2. Mismatched workers with the largest penalties of nearly 30% or more are mismatched because of working conditions, job location, family responsibilities, and a job not being available in their educational field. However, when workers are mismatched due to pay or promotion or a career change, they incur much smaller earnings penalties, not much different than those who are moderately mismatched.

(Table 2 here)

Decomposing the mismatch penalty

Given that the mismatch penalty is large and statistically significant, we then examine the proportion of the differential that is explained versus unexplained. That is, we attempt to answer the question: can the mismatch penalty be explained by observed characteristics such as education and occupation field or is it the unobserved characteristics, captured in the estimated coefficients, which drive the gap? Thus, we use an Oaxaca-style decomposition to decompose the wage differential two ways: between the matched and the moderately mismatched versus the severely mismatched and between the matched versus the moderately and severely mismatched.

Table 3 summarizes the Oaxaca-style decomposition results. Comparing matched and moderately mismatched workers to severely mismatched workers in the top panel, we find that observed characteristics explain 30% of the mismatch penalty. These measurable characteristics explain more of the penalty (34%) for men, and less of the penalty (25%) for women. Comparing matched workers to moderately and severely mismatched workers, the observed characteristics explain only 24% of the penalty, and the gender differences follow the same pattern as the previous decomposition. In general, educational qualifications and experience seem to drive the explained portion of the mismatch penalty. The remaining differential, accounting for about two-thirds to three-quarters of the mismatch penalty, is driven by unobserved characteristics. It is the varying returns to observed characteristics, such as the returns to education and experience that explains the majority of the penalty. These varying returns can be interpreted as involuntary reasons for mismatch (for example, discrimination or a lack of adequate graduate level jobs in the labor market) or the internal barriers that workers put on themselves such as voluntary choices or lack of ambition.

(Table 3 here)

Indeed, some evidence of this can be found in Table 4, where we calculate decompositions between the matched and those mismatched because of pay or because of no job being available where the former reason for mismatch is arguably more 'voluntary' than the latter. The results there suggest that, not only is the differential smaller for those who are severely mismatched due to pay (as expected), but that the 'explained' part of the decomposition is relatively higher for this group of workers. For those whose mismatch is attributed to no job being available, worker characteristics do not explain much of the differential (no more than 20%).

(Table 4 here)

The earnings distribution and the mismatch penalty

Most empirical research has focused on OLS measures of the penalty, essentially looking at mean effects. We extend the research to examine the penalty across the earnings distribution. The first step is to decompose the difference in earnings between the matched and mismatched into a difference in means and a difference in distributions. Following Bender (2003) and Belman and Heywood (2004), we employ a mean squared error (MSE) decomposition (see Appendix 1 for a brief sketch of the methodology). Similar to the Oaxaca decompositions, this method decomposes earnings differentials for two groups: between the matched and moderately mismatched workers to severely mismatched workers, and matched workers to moderately and severely mismatched workers. Table 5 contains the results of these two decompositions. Generally, the results suggest a relatively large difference in distributions that is between 49 and 58% of total MSE. Small variations are evident when the sample is divided by gender. In both comparisons, the total MSE differential is larger for men relative to women. Furthermore, it is the difference in distributions that explains a larger percentage of the differential for men compared to women. For example, when matched and moderately mismatched workers are compared to severely mismatched workers, the difference in distributions explains 55% of the differential, relative to 47% for women. When matched workers are compared to moderately and severely mismatched workers, the difference in distributions explains 62% of the differential for men and 57% of the differential for women. Thus, the results of this decomposition clearly point to a need to examine earnings differentials along the earnings distribution.

(Table 5 here)

Mismatch penalty using quantile regression

Given that the MSE decompositions indicate that the difference in earnings distribution play a sizeable role in explaining the mismatch penalty, the next step is to estimate the effect of mismatch across the earnings distribution using quantile regression (Buchinsky, 1994, 1998).

Figure 1 illustrates the mismatch penalty across the earnings distribution for all workers. Moderate mismatch imposes a small penalty that is relatively flat, although the gap narrows slightly from 9 to 5% across the earnings distribution. The mismatch for severely mismatched workers is quite different. Regardless of position in the earnings distribution, severely mismatched worker's earnings do not catch up to the earnings of workers who are only moderately mismatched. Severely mismatched workers incur penalties around 30% in the bottom decile, 20% at the median and 13% in the upper decile. Therefore, we conclude that while educational mismatch does come with an earnings penalty and this penalty shrinks as workers earn more (as found in Budria and Moro-Egido, 2008 and McGuinness and Bennett, 2007), it does not dissipate completely, even for the top earners in the sample.

(Figure 1 here)

Next, given that we estimate small gender differences in the MSE results, we split the sample by gender in the quantile regressions to compare the mismatch penalties across the male and female earnings distributions. In general, the patterns are consistent across genders - see Figures 2 and 3. Two exceptions are noted. First, among moderately mismatched workers in the upper three quartiles, females incur a penalty that is about one percentage point larger compared to males. Second, among severely mismatched workers in the upper decile, females continue to reduce their penalty on the same trajectory, whereas the penalty among males increases from 12 to 19%.

(Figures 2 and 3 here)

Finally, we draw upon the NSCG's question that asks severely mismatched workers what is the most important reason they work in a field that is not at all related to their educational field. Quantile regression is applied to six reasons for mismatch: pay and promotion opportunities, working conditions, job location, change in career or professional interests, family-related reasons, and job not available. Figure 4 shows the mismatch penalties across the earnings distribution for each reason. The penalties follow the same general shape, which increase across the distribution with a small U-shape (widening of the penalty) in the second decile. However, the sizes of the penalties indicate considerable differences in penalties according to different reasons for mismatch. Unsurprisingly, when workers are mismatched due to pay and promotion opportunities, they experience the smallest penalties that range from about 2 to 18%. Some workers in top decile come close to experiencing no penalty at all compared to matched workers. Changes in career are also associated with smaller penalties that range from about 7 to 22%.

(Figure 4 here)

Workers who are mismatched due to a change in career, family-related reasons, location, and job availability incur the largest penalties. Some interesting findings by reason follow. When workers are mismatched due to family-related reasons, they incur penalties from about 28 to 34% in the bottom half of the earnings distribution, however the penalty narrows significantly in the upper half of the distribution. These workers experience an estimated 22% penalty in the 90th percentile on a 10% penalty in the 99th percentile. When workers are mismatched due to location and job availability, their penalties widen in the top decile, increasing from 24 to 32% and 26 to 37%, respectively.

Again, we split the sample by gender, this time investigating gender differences by the reason for mismatch across the earnings distribution. The results do indicate some differentials by gender, but the differences are not often statistically significant (results available from the authors).

IV. CONCLUSION

While there has been a good deal of research on the earnings penalty for educational mismatch, this has primarily examined differences at the mean level of earnings. This paper is one of the few that have extended the analysis to examining the penalties across the earnings distribution and the first to decompose the earnings differentials for mismatch penalties into 'explained' and 'unexplained' components, to examine whether incomparability in earnings is mostly due to differences on average or in the distribution and to identify whether the reason for mismatch impacts the penalty across the distribution.

Using data from a large, nationally representative dataset of highly educated workers in the US, we find that worker characteristics explain a relatively small proportion of the earnings differentials, particularly for involuntary types of mismatch. In addition, incomparability in earnings between the matched and mismatched is primarily a function of differences in the distributions, indicating the need to focus more on the impact of mismatch throughout the distribution. Quantile regressions suggest particularly large mismatch penalties among the low paid and for workers who are mismatched due to no job being available, to family or to locational reasons.

As there has been little research done on the distributional aspects of mismatch, the results from this paper suggest a number of interesting avenues for future research. For example, the sample of highly educated workers in STEM fields is quite selected and so it would be interesting to see if the results would be replicated in a more broadly representative sample of workers.

Individual heterogeneity could also play a part in the results and the use of panel estimation would be an interesting extension. Finally, this research has only focused on the labor supply side of the market. Bringing in the labor demand side by including firm characteristics would add an interesting dimension to the research.

REFERENCES

Arias, O., K. Hallock, and W. Sosa-Escuderoz. 2000. “Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression using Twins Data.” *Empirical Economics* 26(1), 7-40.

Baker, J.G., R.K. Craft, and M.G. Finn. 2010. “Advantages and disadvantages: job satisfaction of science and engineering baccalaureates completing doctoral verses professional degrees.” Southern Utah University Department of Economics Working Paper.

Belman, D. and J. Heywood. 1997. “Sheepskin Effects by Cohort: Implications of Job Matching in a Signaling Model.” *Oxford Economic Papers* 49(4), 623-37.

Bender, K.A. 2003. “Examining equality between public- and private-sector wage distributions.” *Economic Inquiry* 41(1), 62-79.

Bender, K.A. and J.S. Heywood. 2006. “Job satisfaction of the highly educated: The role of gender, academic tenure, and comparison income.” *Scottish Journal of Political Economy* 53(2), 253-79.

Bender, K.A. and J.S. Heywood. 2009. “Educational mismatch among Ph.D.s: Determinants and consequences.” In *Science and engineering careers in the United States: An analysis of markets and employment*, ed. by R.B. Freeman and D.F. Goroff, Chicago, IL: University of Chicago Press and National Bureau of Economic Research, 229-55.

Bender, K.A. and K. Roche. 2013. “Educational mismatch and self-employment.” *Economics of Education Review* 34, 85-95.

Borghans, L. and A. de Grip (eds). 2000. *The overeducated worker*, Cheltenham, UK: Edward Elgar.

Buchinsky, M. 1994. "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression." *Econometrica* 62(2), 405-58.

Buchinsky, M. 1998. “Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research.” *The Journal of Human Resources* 33(1), 88-126.

Budría, S. 2011. “Are Educational Mismatches Responsible for the ‘Inequality Increasing Effect’ of Education?” *Social Indicators Research* 102, 409-437.

Budría, S. and A. Moro-Egido. 2008. “Education, Educational mismatch and Wage Inequality: Evidence for Spain.” *Economics of Education Review* 27(3), 332-341.

Chevalier, A. 2003. “Measuring over-education.” *Economica* 70, 509-31.

Hernández, L. and L. Serrano. 2012. “Overeducation and its effects on wages: a closer look at the Spanish regions.” *Investigaciones Regionales* 24, 59-90.

Groot, W. and H.M. van den Brink. 2000. Over-education in the labor market: A meta-analysis," *Economics of Education Review* 19, 149-58.

McGuinness, S., and J. Bennett. 2007. "Overeducation in the graduate labour market: A quantile regression approach." *Economics of Education Review* 26, 521-531.

Ordine, P. and R. Giuseppe. 2015. "The effect of family background, university quality and educational mismatch on wage: An analysis using a young cohort of Italian graduates." *Education Economics* 23(2), 213-237.

Wolbers, M. 2003. "Job mismatches and their labor-market effects among school-leavers in Europe." *European Sociological Review* 19, 249-66.

Yu,K., Z. Lu, and J. Stander. 2003. "Quantile regression: applications and current research areas." *The Statitician* 52(3), 331-350.

Table 1. Rates of mismatch, rates of the reasons of severely mismatched workers, and mean and median earnings by mismatch type

Rates of mismatch	Full	Female	Male
Closely Matched	62.1%	64.1%	60.9%
Moderately Mismatched	23.8	21.3	25.3
Severely Mismatched	14.1	14.7	13.7
Rates of the reasons for severe mismatch			
Pay	31.2%	27.7%	34.7%
Conditions	9.2	10.0	8.6
Location	6.1	6.0	6.3
Career	20.7	20.1	21.0
Family	8.4	13.1	5.6
No job	15.7	15.5	15.8
Other	7.7	7.2	8.0
Mean Annual Pay			
Closely Matched	$35,319	$29,621	$38,772
Moderately Mismatched	31,915	26,811	34,384
Severely Mismatched	26,441	23,593	28,197
Median Annual Pay			
Closely Matched	$30,594	$26,042	$33,654
Moderately Mismatched	28,846	24,038	30,831
Severely Mismatched	22,436	19,872	24,038

Data source: Data are for 66,172 full-time workers from the 2003 NSCG.

Table 2. Selected results from Ordinary Least Squares regression

Variable	Full	Female	Male
Moderately mismatched	-0.068*** (0.004)	-0.076*** (0.007)	-0.064*** (0.006)
Severely mismatched	-0.224*** (0.006)	-0.209*** (0.009)	-0.225*** (0.007)
Reasons for severe mismatch			
Pay and promotion opportunities	-0.083*** (0.004)	-0.064*** (0.015)	-0.094*** (0.011)
Working conditions	-0.271*** (0.016)	-0.246*** (0.023)	-0.273*** (0.021)
Location	-0.339*** (0.019)	-0.321*** (0.029)	-0.340*** (0.024)
Career change	-0.177*** (0.011)	-0.142*** (0.017)	-0.191*** (0.014)
Family-related	-0.332*** (0.016)	-0.333*** (0.020)	-0.282*** (0.025)
No job available	-0.388*** (0.012)	-0.340*** (0.019)	-0.407*** (0.015)
Other reason	-0.317*** (0.017)	-0.291*** (0.027)	-0.323*** (0.021)

Data source: Data are for 66,172 full-time workers from the 2003 NSCG.
Notes: Results are from a log hourly earnings regression. Regressions control for gender, age, age squared, race, marital status, US citizenship, experience, experience squared, tenure, tenure squared, highest degree, occupation, industry, disability status, employment sector, firm size, and region. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table 3. Oaxaca-style decomposition results

	Full	Female	Male
Matched and moderately mismatched vs severely mismatched			
Total differential	0.288***	0.241***	0.309***
	(0.006)	(0.010)	(0.008)
Explained	0.087***	0.060***	0.105***
	(0.004)	(0.006)	(0.005)
% of Differential	30%	25%	34%
Unexplained	0.201***	0.181***	0.204***
	(0.006)	(0.010)	(0.008)
% of Differential	70%	75%	66%
Matched vs moderately and severely mismatched			
Total differential	0.171***	0.159***	0.189***
	(0.004)	(0.004)	(0.005)
Explained	0.041***	0.027***	0.061***
	(0.003)	(0.005)	(0.004)
% of Differential	24%	17%	32%
Unexplained	0.130***	0.132***	0.128***
	(0.004)	(0.007)	(0.006)
% of Differential	76%	83%	68%

Data source: Data are for 66,172 full-time workers from the 2003 NSCG.
Notes: Results are from an Oaxaca decomposition that uses a log hourly earnings regression. Regressions control for gender, age, age squared, race, marital status, US citizenship, experience, experience squared, tenure, tenure squared, highest degree, occupation, industry, disability status, employment sector, firm size, and region. Standard errors are in parentheses. The closely matched group is the reference wage structure. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table 4. Oaxaca-style decomposition results for reasons of pay and job not available

	Full		Female		Male	
	Pay	No Job	Pay	No Job	Pay	No Job
Total differential	0.131***	0.483***	0.074***	0.396***	0.177***	0.536***
	(0.010)	(0.015)	(0.017)	(0.024)	(0.013)	(0.019)
Explained	0.043***	0.072***	0.013	0.039***	0.074***	0.100***
	(0.006)	(0.008)	(0.009)	(0.012)	(0.007)	(0.010)
% of Differential	33%	15%	18%	10%	42%	19%
Unexplained	0.089***	0.411***	0.061***	0.0357***	0.103***	0.436***
	(0.010)	(0.014)	(0.018)	(0.023)	(0.013)	(0.018)
% of Differential	68%	85%	82%	90%	58%	81%

Data source: Data are for 66,172 full-time workers from the 2003 NSCG.
Notes: Results are from an Oaxaca decomposition that uses a log hourly earnings regression. Regressions control for gender, age, age squared, race, marital status, US citizenship, experience, experience squared, tenure, tenure squared, highest degree, occupation, industry, disability status, employment sector, firm size, and region. Standard errors are in parentheses. The closely matched group is the reference wage structure. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

Table 5. Mean Square Error decomposition results

	Full	Female	Male
Matched and moderately matched vs severely mismatched			
Total differential	0.089	0.074	0.103
Difference in means	0.045	0.039	0.046
% of Differential	51%	53%	45%
Difference in distribution	0.044	0.035	0.057
% of Differential	49%	47%	55%
Matched vs moderately and severely mismatched			
Total differential	0.050	0.047	0.050
Difference in means	0.021	0.02	0.019
% of Differential	42%	43%	38%
Difference in distribution	0.029	0.027	0.031
% of Differential	58%	57%	62%

Data source: Data are for 66,172 full-time workers from the 2003 NSCG.
Notes: Results are from a Mean Square Error decomposition that uses a log hourly earnings regression. Regressions control for gender, age, age squared, race, marital status, US citizenship, experience, experience squared, tenure, tenure squared, highest degree, occupation, industry, disability status, employment sector, firm size, and region.

Figure 1. Mismatch penalty at each percentile of the earnings distribution for all workers

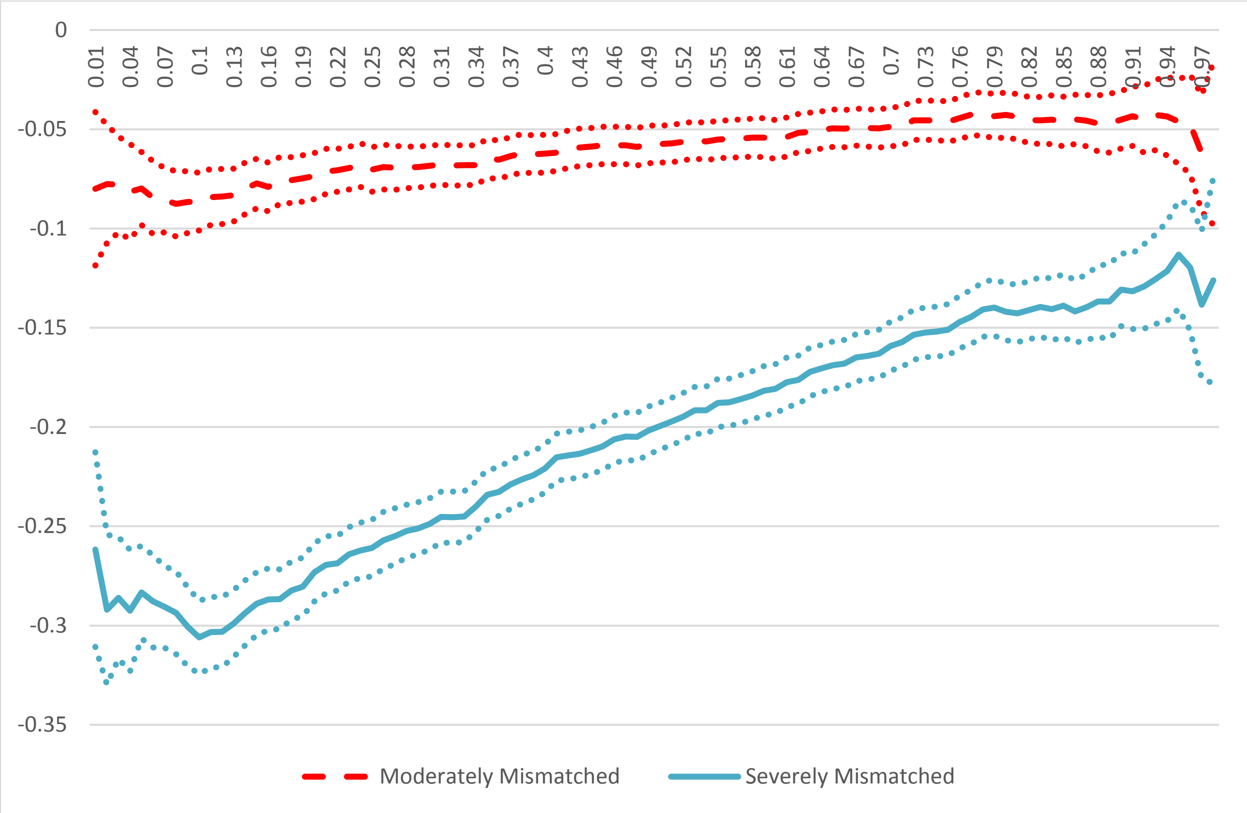


Figure 2. Mismatch penalty at each percentile in the earnings distribution for males

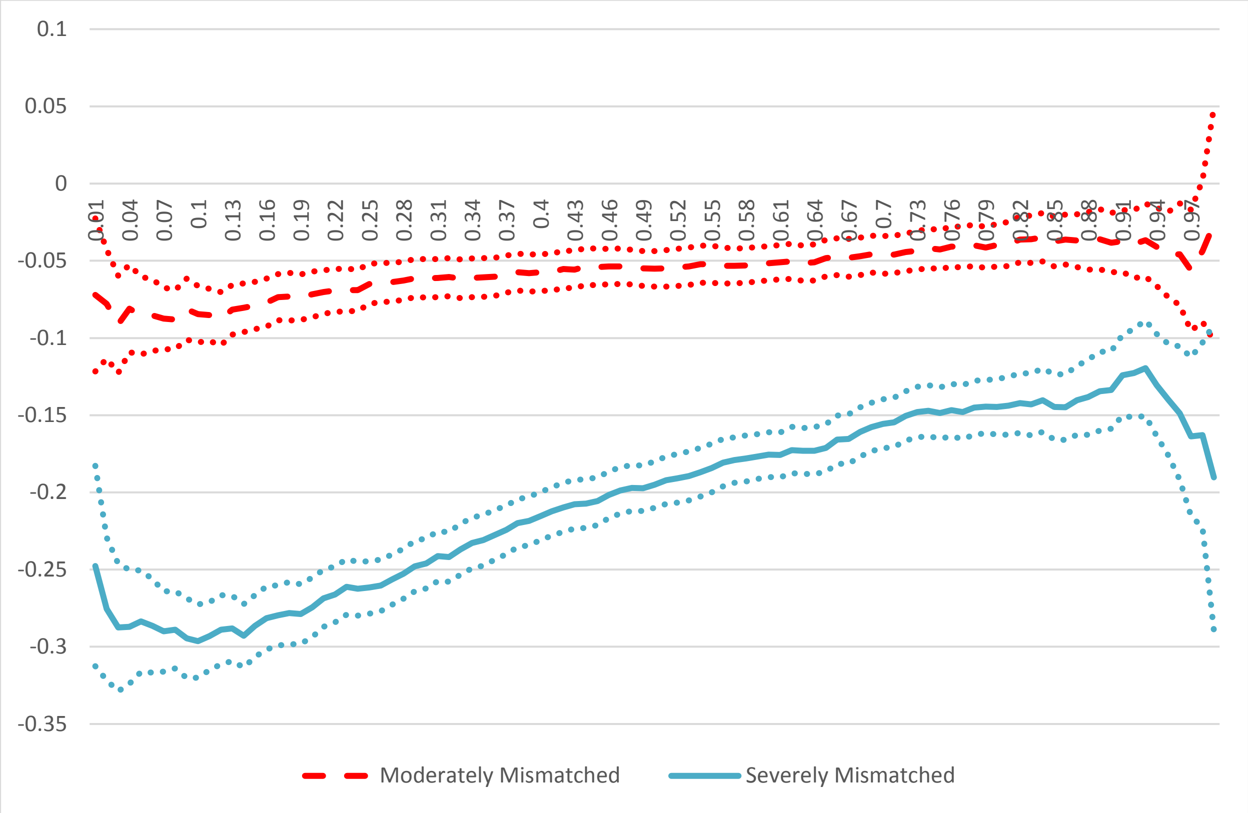


Figure 3. Mismatch penalty at each percentile in the earnings distribution for females

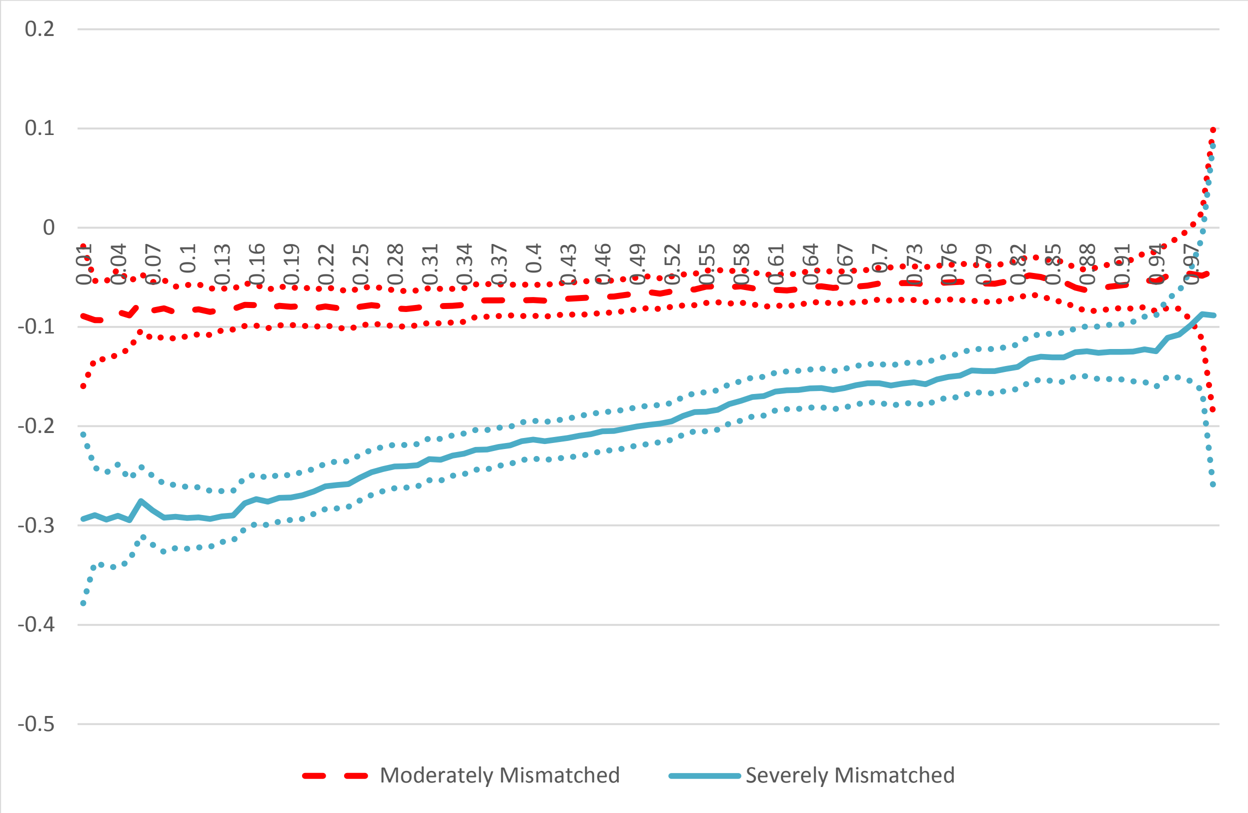
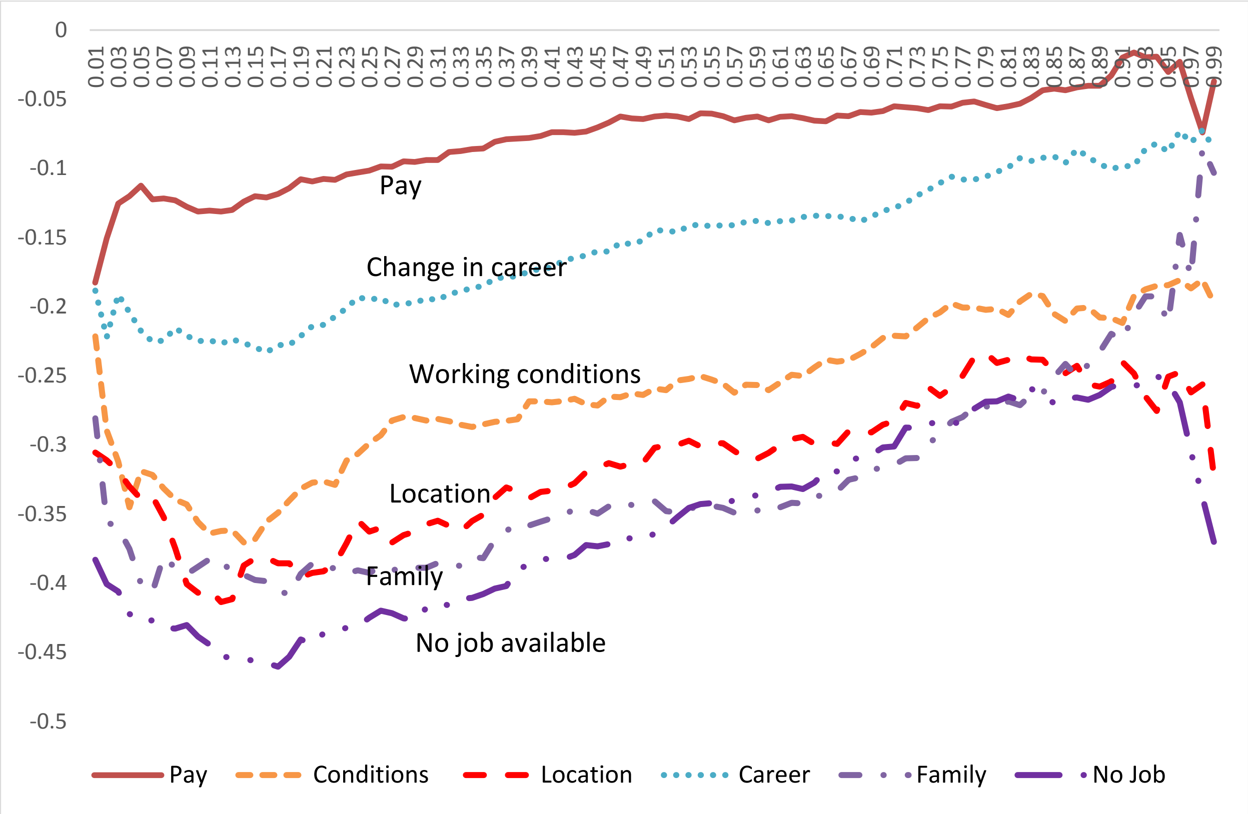


Figure 4. Mismatch penalty by reason at each percentile in the earnings distribution for all workers



Appendix 1. Sketch of Mean Squared Error (MSE) Decomposition Methodology.

While full details of the MSE methodology can be found in Bender (2003) and Belman and Heywood (2004) who applied the methodology to examining differences in public-private earnings differences, below we briefly describe the methodology as applied to the earnings differentials between matched and mismatched workers.

Define the earnings differential for each worker, i, between matched (m) and mismatched (mm):

(A.1) $$\theta_i = ln\widehat{W}_i^{m} - ln\widehat{W}_i^{mm} = X_i\hat{\beta}^{m} - X_i\hat{\beta}^{mm}$$

The definition of the MSE in terms of equation (A.1) for n workers is:

(A.2) $$\text{MSE}(\theta) = \left(\frac{1}{n}\right) * \sum_i (\theta_i - \theta_i^c)$$

Comparability implies that earnings are the same regardless of the match or that $\theta_i^c = 0$. If wages are comparable (that is, $\theta_i^c = 0$), it can be shown that the MSE can be represented as:

(A.3) $$\text{MSE}(\theta) = \left(\frac{1}{n}\right) * \sum_i (\theta_i - \bar{\theta})^2 + \bar{\theta}^2 = var(\theta) + \bar{\theta}^2,$$

where $\bar{\theta}$ is the mean predicted wage over the sample. Equation (A.3) decomposes incomparability into two separate dimensions – differences in the distribution of the predicted differential and the squared average of the predicted differential. If all workers have the same value of θ, then there is no difference in the distribution of predicted wages by mismatch status. This would mean that the comparability of wages across mismatch status is only due to differences in average wages across groups. The second term is zero if there are no differences in average wages across groups and any incomparability is highlighted by differences in distributions. Of course, it is likely that there are differences in both the distributions and average wages, and by decomposing using the MSE method, we can determine the relative importance of the two determinants of the MSE.

Endnotes

[1] Other papers focus on the role of educational mismatch on wage inequality. For example, Budrıa (2011) finds that educational mismatch does not drive the positive effect of education on wage inequality in Portugal and Europe, and Ordine and Rose (2015) argue that educational mismatch does explain some wage inequality among college graduates in Italy.

[2] Due to the ambiguity around the “other” reason for severe mismatch, our remaining analysis does not present results on this particular reason. We focus on the first six reasons as they allow us to draw more meaningful conclusions from them. However, the “other” reason category is included in all regressions using the reasons for mismatch and are available upon request.