Cornell University
ILR School

Cornell University ILR School
DigitalCommons@ILR

Institute for Compensation Studies

Centers, Institutes, Programs

10-17-2010

Immigrant Heterogeneity and the Earnings Distribution in the United Kingdom and United States: New Evidence from a Panel Data Quantile Regression Analysis

Sherrilyn M. Billger

Carlos LaMarche

Follow this and additional works at: http://digitalcommons.ilr.cornell.edu/ics

This Article is brought to you for free and open access by the Centers, Institutes, Programs at DigitalCommons@ILR. It has been accepted for inclusion in Institute for Compensation Studies by an authorized administrator of DigitalCommons@ILR. For more information, please contact jdd10@cornell.edu.

Immigrant Heterogeneity and the Earnings Distribution in the United Kingdom and United States: New Evidence from a Panel Data Quantile Regression Analysis

Abstract
In this paper we use a relatively new panel data quantile regression technique to examine native-immigrant earnings differentials 1) throughout the conditional wage distribution, and 2) controlling for individual heterogeneity. No previous papers have simultaneously considered these factors. We focus on both women and men, using longitudinal data from the PSID and the BHPS. We show that country of origin, country of residence, and gender are all important deter- minants of the earnings differential. For instance, a large wage penalty occurs in the U.S. among female immigrants from non-English speaking countries, and the penalty is most negative among the lowest (conditional) wages. On the other hand, women in Britain experience hardly any immigrant-native wage differential. We find evidence suggesting that immigrant men in the U.S. and the U.K. earn lower wages, but the most significant results are found for British workers emigrating from non-English speaking countries. The various differentials we report in this paper reveal the value of combining quantile regression with controls for individual heterogeneity in better understanding immigrant wage effects.

Keywords
Immigrants, Earnings, Quantile regression, Panel data

Comments

Suggested Citation
Billger, S. & LaMarche, C. (2010). *Immigrant heterogeneity and the earnings distribution in the United Kingdom and United States: new evidence from a panel data quantile regression analysis* (ICS 2010-002). Retrieved [insert date] from Cornell University, ILR School, Institute for Compensation Studies site: http://digitalcommons.ilr.cornell.edu/ics/2

This article is available at DigitalCommons@ILR: http://digitalcommons.ilr.cornell.edu/ics/2

IMMIGRANT HETEROGENEITY AND THE EARNINGS DISTRIBUTION IN THE UNITED KINGDOM AND UNITED STATES: NEW EVIDENCE FROM A PANEL DATA QUANTILE REGRESSION ANALYSIS*

SHERRILYN M. BILLGER and CARLOS LAMARCHE†

Abstract. In this paper we use a relatively new panel data quantile regression technique to examine native-immigrant earnings differentials 1) throughout the conditional wage distribution, and 2) controlling for individual heterogeneity. No previous papers have simultaneously considered these factors. We focus on both women and men, using longitudinal data from the PSID and the BHPS. We show that country of origin, country of residence, and gender are all important determinants of the earnings differential. For instance, a large wage penalty occurs in the U.S. among female immigrants from non-English speaking countries, and the penalty is most negative among the lowest (conditional) wages. On the other hand, women in Britain experience hardly any immigrant-native wage differential. We find evidence suggesting that immigrant men in the U.S. and the U.K. earn lower wages, but the most significant results are found for British workers emigrating from non-English speaking countries. The various differentials we report in this paper reveal the value of combining quantile regression with controls for individual heterogeneity in better understanding immigrant wage effects.

Keywords: Immigrants; Earnings; Quantile regression; Panel data.

JEL Codes: J31, J61, C21, C23.

*This version: October 17, 2010. First version: March 1, 2010. We are grateful to seminar participants at the seventh IZA annual migration meeting for useful comments.

†Sherrilyn M. Billger: Department of Economics, Illinois State University. Campus Box 4200, Normal, IL 61790. Tel.: +1 309 438 8720. Email: smbillg@ilstu.edu.

Carlos Lamarche: Department of Economics, University of Oklahoma. 202 Hester Hall, 729 Elm Avenue, Norman, OK 73019. Tel.: +1 405 325 5857. Email: lamarche@ou.edu

1. Introduction

It is well-known that immigrant earnings often differ significantly from native earnings. Numerous studies highlight the sources of these earnings differentials, noting that human capital and language skills are particularly important. Thus far, empirical work in this area has used relatively basic regression techniques, including panel data models and quantile regression. From some of these studies we know that individual heterogeneity is an important factor in estimating earnings differentials for immigrants. Furthermore, it behooves us to learn about the effect of immigrant status throughout the conditional wage distribution, not simply at the mean. For instance, ongoing debate about Mexican immigrants to the U.S. can be informed by our results. We expect that workers who immigrate to the U.S. from non-English speaking countries earn lower wages than natives. However, it is entirely possible that the negative wage effect is larger in the left tail of the conditional earnings distribution, suggesting that low-skilled immigrants earn particularly low wages relative to natives. With these issues in mind, we use a new panel data regression technique to simultaneously account for individual heterogeneity and to generate parameter estimates across the conditional distribution of earnings.

We focus this analysis on data from the United States and the United Kingdom, two countries with sizable immigrant populations and clear value for English language ability. Most previous papers in the immigration literature use Census data, and we contribute to this area using two longitudinal studies with substantial immigrant subsamples: the U.S. Panel Study of Income Dynamics and the British Household Panel Survey. The extensive longitudinal data enable us to create specifications similar to those in previous papers, while employing a panel data - quantile regression approach (Koenker 2004, Lamarche 2010) in order to determine better estimates of native-immigrant earnings differentials. These methods have been recently employed in Kniesner, Viscusi and Ziliak (2010) and Kniesner, and Leeth (2010). Furthermore, despite the fact that increasing proportions of women are in the labor force, most previous work focuses only on men, but in this paper we study wages for both men and women.

In this paper we examine immigrant-native wage differentials that differ in many dimensions from the existing work in this area. We simultaneously consider differentials throughout the conditional earnings distribution and the role of unobserved heterogeneity in these differentials. We document substantial differences by gender, country of origin, and destination country. For instance, in Britain, it appears that individual-specific characteristics like language and ability explain much of the immigrant-native differentials across the conditional earnings distribution. We also find that while immigrant men from non-English speaking countries earn significantly less than their native counterparts in the U.K., immigrant women do not seem to experience any significant wage differential. Regardless of origin country, immigrants to the U.S. appear to earn much less than natives. For men, we see sizable differences across the conditional quantiles of earnings, but again, these differences appear largely correlated with individual heterogeneity. Perhaps most importantly, pooled cross-sectional estimates of immigrant-native differentials appear to greatly overstate the actual wage penalties suffered by the lowest earning male immigrants.

A broad segment of the existing literature highlights wage differences between immigrants and natives, often citing the specific sources for these differentials.[1] One source of native-immigrant differentials is whether human capital was obtained in the source or destination country (Friedberg 2000, Bratsberg and Ragan 2002). Occupation is also important (Chiswick and Miller 2007, Chiswick and Miller 2009, Green 1999). Another source of variation within immigrant earnings is race. Clark and Lindley (2006) use a pooled cross-section from the UK Labour Force Survey and find that nonwhite immigrants receive lower pay than white immigrants to the UK.

More recent studies use longitudinal data to investigate the role of language fluency and to determine whether wage gaps persist over time.[2] Chiswick, Lee, and Miller (2005) use longitudinal data from Australia and document the importance of skill and assimilation in improving earnings. The authors contend that longitudinal estimates (generated by an inertia model) mirror those for the cross-section. However, it is

[1]Borjas (1989) provides one of the seminal papers in this area, documenting the effects of assimilation and cohort on immigrant earnings. With data on immigrants to Israel, Weiss, Sauer, and Gotlibovski (2003) find that lifetime earnings are 57% lower than those for natives.

[2]Over the 1960-1992 period, Duleep and Dowhan (2002) find that immigrants experienced more wage growth than natives did.

possible that a panel longer than 3.5 years could generate differences. Lubotsky (2007) uses longitudinal social security data to explore the immigrant/native wage gap. Comparing panel data estimates to those from cross-sectional data, Lubotsky finds that the gap is more persistent, suggesting that cross-sectional estimates (such as those from the Census) are biased. Furthermore, his results reveal that assimilation is less prevalent in panel data analyses, due in part to likely outmigration of some in immigrant cohorts.

In a unique study that investigates both genders, Butcher and DiNardo (2002) compare wage density estimates using Census PUMS data, and find important differences for men and women.[3] Bleakley and Chin (2004) contribute to the literature on language skills by employing IV techniques. They highlight that the time of arrival has an important impact on learning the language of the destination country. By completing education in the US, adults who arrived as children see significant returns to their language proficiency.[4] In another panel data study, Hum and Simpson (2004) report that omitted variables bias immigration analyses. Using fixed effects and IV techniques, the authors document lasting immigrant/native wage differentials, providing a contrast to other studies that report assimilation over time.

Dustmann and Van Soest (2002) provide a valuable contribution to the literature by investigating unobserved heterogeneity and measurement error. Using German panel data and minimum distance estimation, they report pronounced downward bias in the returns to language skill. Dustmann and Van Soest provide evidence that estimates presented in the literature are biased, and that language is even more important in determining immigrant earnings. Finally, Chiswick, Le, and Miller (2008) provide one of few immigration studies that explore the distribution of returns to education and experience. Using quantile regression and data from the US and the Australian Censuses, they find that schooling and experience serve to expand the earnings distribution; i.e., higher returns occur at higher values in the conditional earnings distribution. The authors also find that the adult male native/immigrant wage differential is higher at higher quantiles.

[3]Their results also support Lubotsky's (2007) finding that changes in the returns to skill contribute substantially to the growth in the gap between immigrants and natives.

[4]See also Chiswick and Miller (1995).

2. Quantile Regression Models and Methods

Quantile regression is a robust estimation approach that enables the estimation of native-born/immigrant earnings differentials at different quantiles of the conditional earnings distribution. This approach was recently used (in the cross-section) in Chiswick, Le, and Miller (2008). In this paper, we present a newly developed approach that not only allows us to explore these differentials, but also controls for native-born and immigrant unobserved heterogeneity.

We begin with a model of earnings given by,

$$w_{it} = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \alpha_i + u_{it}, \tag{2.1}$$

where w_{it} denotes the logarithm of earnings for individual i at time t, $\boldsymbol{x}_{it}$ is a $p \times 1$ vector of independent variables that includes an intercept, α_i is a latent term denoting individual-specific differences in language ability and skills, and u_{it} is an independent and identically distributed (iid) random variable with distribution F. We assume that u_{it} is independent of $\boldsymbol{x}_{it}$.[5]

The latent individual variable may well be correlated with the independent variables, so it is natural to estimate a conditional mean model with fixed effects,

$$\mathbb{E}(w_{it}|\boldsymbol{x}_{it}, \alpha_i) = \boldsymbol{x}_{it}'\boldsymbol{\beta} + \alpha_i.$$

Our approach here builds upon this classical empirical strategy, estimating a fixed effects version of the conditional quantile regression model,

$$Q_{W_{it}}(\tau_j|\boldsymbol{x}_{it}, \alpha_i) = \boldsymbol{x}_{it}'\boldsymbol{\beta}(\tau_j) + \alpha_i, \tag{2.2}$$

where τ_j is a quantile in $(0, 1)$ and $Q(\cdot|\cdot)$ is the τ_j-th conditional quantile function. The parameter $\boldsymbol{\beta}(\tau_j)$ provides an opportunity to investigate how the independent variables influence the location, scale and shape of the conditional distribution of earnings. For instance, if we have an iid error term distributed as F and one covariate, the quantile functions $Q_{W_{it}}(\tau_j|\boldsymbol{x}_{it}, \alpha_i)$ are parallel lines with parameter $(\beta_0(\tau_j), \beta_1))'$, and

[5]Under similar conditions, Graham, Hahn, and Powell (2009) show that there is no incidental parameter problem in a non-differentiable panel data model. Recent approaches investigate the estimation of models under different assumptions. Harding and Lamarche (2009) estimate a model with endogenous covariates and Galvao (2009) analyses the fixed effects estimator introduced by Koenker (2004) in a dynamic framework.

the distance between them is the difference between the intercept terms. The model also includes an individual effect α_i. The individual effect represents a pure location shift effect on the conditional quantiles of earnings, implying that the conditional distribution for each individual has the same shape. The locations of these lines differs, as long as language ability and skills within α_i are different. Notice that the individual effect does not represent a distributional shift.[6]

2.1. **Panel Data Estimators.** We estimate the conditional quantile earnings function (2.2) employing the class of panel data estimators introduced by Koenker (2004) and Lamarche (2010). Koenker's (2004) penalized quantile regression estimator can be obtained as the solution of a problem similar to,

$$\min_{\beta,\alpha\in\mathcal{B}\times\mathcal{A}} \sum_{j=1}^{J}\sum_{t=1}^{T}\sum_{i=1}^{N} \omega_j \rho_{\tau_j}(w_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}(\tau_j) - \alpha_i) + \lambda Pen(\alpha) \tag{2.3}$$

where $\rho_{\tau_j}(u) = u(\tau_j - I(u \leq 0))$ is the standard quantile loss function (see, e.g., Koenker 2005), ω_j is a relative weight given to the j-th quantile, and λ is the tuning parameter.

This method proposes to jointly estimate the parameter $\boldsymbol{\beta}$ and a vector of N individual effects $\boldsymbol{\alpha}$, because the standard panel data transformations are not available in quantile regression. This strategy increases the variability of the estimator of the parameter of interest $\boldsymbol{\beta}$. To attenuate this variability, we minimize over a weighted sum of quantile check functions including an additional (penalty) term $\lambda Pen(\alpha)$. This penalty term shrinks the individual effects toward zero and the degree of shrinkage is controlled by λ. For $\lambda = 0$, we have the fixed effects estimator, while for $\lambda > 0$, we have the penalized estimator with fixed effects. As λ approaches infinity, this becomes a pooled estimator.[7]

2.1.1. *Pooled and fixed effects methods.* A tentative approach for estimating native-born/immigrant differentials across the conditional distribution of earnings is Koenker and Bassett's (1978) method applied to the pooled data. This estimator is equivalent to what arises from (2.3) by letting the shrinkage parameter $\lambda \to \infty$. The method

[6]It may be unrealistic to estimate it when the number of observations on each individual is small (Koenker 2004).

[7]For additional methodological details, please see Koenker (2004, 2005) and Lamarche (2010).

gives the opportunity of estimating the standard quantile regression model without fixed effects, $Q_{W_{it}}(\tau|\boldsymbol{x}_{it}) = \boldsymbol{x}_{it}'\boldsymbol{\beta}(\tau)$. The quantile regression estimator for the pooled data is defined as,

$$\hat{\boldsymbol{\beta}}(\tau) = \underset{\beta\in\mathcal{B}}{\operatorname{argmin}} \sum_{t=1}^{T}\sum_{i=1}^{N} \rho_\tau(w_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}). \tag{2.4}$$

This method is convenient because it allows one to estimate time-invariant effects of interest associated with native-born/immigrant earnings differentials. However, as in the least squares case, omitting relevant variables may generate biases. Under the conditions of equation (2.1), we may briefly derive the omitted variable bias in our quantile regression model, which can be obtained as a direct application of Angrist, Chernozhukov, and Fernandez-Val (2006), Theorem 2. Let $\boldsymbol{\beta}_s(\tau)$ be the slope coefficient in a quantile regression of earnings on $\boldsymbol{x}$ and $\boldsymbol{\beta}_l(\tau)$ the regression coefficient of earnings on a vector indicating individual effects, $\boldsymbol{z}$, and $\boldsymbol{x}$. The relation between these coefficients can be written as,

$$\boldsymbol{\beta}_s(\tau) = \boldsymbol{\beta}_l(\tau) + (\mathbb{E}(\omega_\tau \boldsymbol{x}\boldsymbol{x}'))^{-1}\mathbb{E}(\omega_\tau \boldsymbol{x}\boldsymbol{z}\alpha),$$

where $\omega_\tau = \int_0^1 f_{e_\tau}(u\Delta_\tau|\boldsymbol{x},\boldsymbol{z})du/2$, $\Delta_\tau = \boldsymbol{x}'\boldsymbol{\beta}_s(\tau) - Q_W(\tau|\boldsymbol{x},\alpha)$, and $e_\tau = W - Q_W(\tau|\boldsymbol{x},\alpha)$. It is immediately apparent that the pooled quantile regression estimator $\hat{\boldsymbol{\beta}}_s(\tau)$ can be biased if $\boldsymbol{x}$ and α are not independent.

Although standard panel data methods offer the possibility of estimating conditional mean models while controlling for individual heterogeneity, until recently, few papers have estimated conditional quantile function with individual specific effects. It is possible to control for unobserved heterogeneity by introducing individual effects as in Koenker (2004),

$$(\hat{\boldsymbol{\beta}}(\tau)', \hat{\boldsymbol{\alpha}}')' = \underset{\beta,\alpha\in\mathcal{B}\times\mathcal{A}}{\operatorname{argmin}} \sum_{j=1}^{J}\sum_{t=1}^{T}\sum_{i=1}^{N} \rho_{\tau_j}(w_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}(\tau_j) - \alpha_i). \tag{2.5}$$

This method allows us to address the possibility of endogenous covariates, but its use is limited if we are most interested in time-invariant native-born/immigrant earnings differentials.

2.1.2. *Penalized quantile regression.* The penalized estimator with fixed effects enables us to estimate these differentials while controlling for individual-specific factors such as language ability and skills. The estimator is defined as,

$$(2.6)\quad (\hat{\boldsymbol{\beta}}(\tau,\lambda)',\hat{\boldsymbol{\alpha}}(\lambda)')' = \underset{\beta,\alpha\in\mathcal{B}\times\mathcal{A}}{\operatorname{argmin}} \sum_{j=1}^{J}\sum_{t=1}^{T}\sum_{i=1}^{N} \rho_{\tau_j}(w_{it} - \boldsymbol{x}_{it}'\boldsymbol{\beta}(\tau_j) - \alpha_i) + \lambda \sum_{i=1}^{N} \rho_{0.5}(\alpha_i).$$

It is immediately apparent that for obtaining $\hat{\boldsymbol{\beta}}(\boldsymbol{\tau},\lambda)$, we need to select the quantiles τ and the parameter λ. As is standard in the quantile regression literature, we consider five quantiles $\tau = \{0.1, 0.25, 0.5, 0.75, 0.9\}$ to investigate the covariate effect across the conditional distribution of the response variable.

The selection of λ is of fundamental importance. This choice should be made to reduce the additional variability introduced by the estimation of the individual effects, as mentioned above. Following Lamarche (2010), we select λ considering a simple variance minimizing strategy:

$$(2.7)\qquad \hat{\lambda} = \arg\inf_{\lambda} \{tr\boldsymbol{\Sigma}\},$$

where $tr\boldsymbol{\Sigma}$ is the trace of the asymptotic covariance matrix. This covariance matrix is obtained by a bootstrap procedure. The basic strategy accommodates to forms of heteroscedasticity replacing pairs $\{(\boldsymbol{w}_i, \boldsymbol{x}_i) : i = 1, \ldots, N\}$ over cross-sectional units i. We draw a worker from a sample of natives and immigrant workers and we include all T observations for that worker. We continue sampling workers (with replacement) as indicated before until we obtain a sample of N cross-sectional units. Using this new sample, we compute penalized estimates $\{\boldsymbol{\beta}^*(\boldsymbol{\tau},\lambda), \boldsymbol{\alpha}^*(\lambda)\}$ for a given value of λ. We reiterate this procedure to obtain the standard error of the estimator.

Alternatively, we consider a method that is similar to the classical cross validation approaches (e.g., CV, GCV). We tentatively select the tuning parameter λ following a procedure motivated by the standard AIC-type approach, $\hat{\lambda} = \arg\inf \|\hat{u}(\tau,\lambda)\|_1 + \mathrm{df}_\lambda/(2NT)$, where $\hat{u}(\tau,\lambda) = w - \boldsymbol{x}'\hat{\boldsymbol{\beta}}(\tau,\lambda) - \hat{\boldsymbol{\alpha}}(\lambda)$ and df_λ is the number of nonzero estimated parameters. The number of nonzero estimated coefficients represents a simple estimate of the degrees of freedom. This λ selection device is time consuming and needs to be implemented by considering a grid.

3. Data

In this analysis we use samples from the U.K. and the U.S. The U.K. data are from the British Household Panel Survey (BHPS), and our sample of 2255 natives and immigrants is from 1991 through 2002. The BHPS provides gross weekly earnings for all employed individuals, and we convert these amounts to real 2002 pounds using the U.K. Consumer Prices Index. We would prefer to analyze hourly wages, and made a simple conversion from weekly to hourly using reported hours worked. Unfortunately, such a conversion yields spurious correlation and measurement error, so we focus on weekly wages instead. The U.S. data are from the Panel Study of Income Dynamics (PSID) that is administered on odd years. Our sample of 3676 workers begins in 1997 and continues every other year through 2005. Hourly earnings are converted to real 2005 dollars using the Consumer Price Index for all urban consumers.

Descriptive statistics for our samples are listed in Table 3.1, separately by gender and immigrant status. Average weekly wages in the U.K. are £459 and £255 for natives, and £542 and £304 for immigrants. In the U.S., native men earn $26 per hour on average, and women earn $18, compared to $15 for immigrants. The gender wage gap appears greater in the U.K., but it is important to remember that the British data are reported weekly, not hourly. Furthermore, immigrants in the U.K. have higher average earnings than their native counterparts, while the opposite is true in the U.S.

To aid comparison across the BHPS and PSID surveys, we attempt to match education variables. Intermediate qualifications involve a roughly equivalent amount of study time in the U.S. and U.K., for example, yielding a high school diploma through BA degree in the US or O level through 1st Degree in the U.K. Among natives, 85% of men and 84% of women have this level of education in Britain, and 81% of men and 85% of women have this in the States. Advanced qualifications are defined as professional and doctoral degrees in the U.S. or Higher Degree in the U.K. Native-born Americans have achieved this level of education at the rates of 11% of men and 10% of women, while Britains have attained at 5% of men and 2% of women. Immigrants are comparably well-educated in the U.K., and males have more schooling than their native peers; 13% of men and 9% of women have advanced qualifications.

	BHPS		PSID	
Variables	Males	Females	Males	Females
	Native-born workers			
Wage rate	459.246	255.306	26.450	18.296
	(255.905)	(184.795)	(27.442)	(14.860)
Intermediate qualifications	0.854	0.842	0.813	0.848
	(0.354)	(0.365)	(0.390)	(0.359)
Advanced qualifications	0.054	0.024	0.112	0.102
	(0.226)	(0.154)	(0.316)	(0.303)
Experience	20.676	22.311	24.363	24.546
	(8.99)	(9.658)	(8.670)	(8.482)
Married	0.676	0.686	0.844	0.705
	(0.468)	(0.464)	(0.363)	(0.456)
Hours	39.118	29.378	45.966	39.035
	(6.881)	(10.845)	(10.360)	(10.640)
Union	0.432	0.448	0.200	0.159
	(0.495)	(0.497)	(0.400)	(0.366)
	Immigrant workers			
Wage rate	541.532	303.991	15.278	14.898
	(385.899)	(230.052)	(12.683)	(13.180)
Intermediate qualifications	0.873	0.799	0.449	0.632
	(0.333)	(0.401)	(0.498)	(0.483)
Advanced qualifications	0.127	0.090	0.087	0.088
	(0.333)	(0.286)	(0.281)	(0.284)
Experience	18.491	22.058	27.050	24.865
	(8.867)	(9.089)	(9.635)	(8.451)
Married	0.770	0.606	0.872	0.668
	(0.421)	(0.489)	(0.334)	(0.472)
Hours	39.343	32.530	46.074	40.015
	(6.068)	(9.650)	(11.773)	(9.674)
Union	0.351	0.480	0.148	0.135
	(0.478)	(0.500)	(0.355)	(0.343)
Non-English Speaking Foreign-Born	0.651	0.740	0.937	0.824
	(0.477)	(0.439)	(0.243)	(0.382)
English Speaking Foreign-Born	0.349	0.260	0.063	0.176
	(0.477)	(0.439)	(0.243)	(0.382)
Years since migration	27.946	26.423	17.433	16.779
	(9.805)	(10.138)	(7.558)	(7.234)
Number of workers	1004	1251	1918	1758
Number of observations	8366	10373	10090	8790

TABLE 3.1. *Descriptive statistics for the samples from the U.K. and U.S. We present descriptive statistics by gender and immigrant status. The data are from British Household Panel Survey (BHPS) and Panel Study of Income Dynamics (PSID).*

Unsurprisingly, immigrants to America have substantially lower levels of education, with only 45% of men and 63% of women holding intermediate qualifications.

Average experience levels are roughly equivalent across the surveys, though the women in our sample have slightly higher (potential) job tenure, as measured by age minus education minus 5. Marital rates are quite high in our sample, with a clear majority in all subgroups. In each subcategory, women work fewer hours per week than men do, and Americans work more hours than Britons. It also appears that immigrants work more hours on average than natives do. We know that unionization is more prevalent in the U.K. than in the U.S. However, in Britain women are more likely than men to be in union jobs. Immigrants comprise a small proportion of our samples, with quite equal gender divisions in the UK, but more men in the US. We also see that the vast majority of immigrants to both countries originate from non-English speaking nations. Finally, within our samples, U.K. immigrants have been in the country for more than 25 years on average, while years since migration is about 17 in the U.S.

4. Empirical Results

This section presents results from two basic specifications. We first estimate a model similar to the one proposed in Chiswick, Le, and Miller (2008), and we mainly focus on the differences between classical quantile regression estimates and panel data quantile regression estimates. The results indicate that unobserved heterogeneity, possibly associated with language ability and skills, plays an important role and needs to be accounted for in earnings regressions for immigrants. We then augment these panel data models with other covariates that are typically considered in the literature.

4.1. **BHPS sample.** Table 4.1 presents quantile regression results obtained from employing the penalized quantile regression approach described in Section 2.1. The upper block of the table shows quantile regression results and the lower block shows penalized quantile regression results. We do not include (fixed effects) results for $\lambda = 0$ because there is no within variation in 'Non-English Speaking Foreign-Born' and 'English Speaking Foreign-Born', two of the main variables of interest in this paper. The last column presents estimates of the classical mean regression model

that is most closely associated with the quantile regression approach. Specifically, the results shown in the upper block correspond to simple OLS and the results shown in the lower block correspond to random effects models.[8]

The first column in Table 4.1 presents our choice for the tuning parameter λ. As discussed before, the selection of the tuning parameter is a fundamental aspect of this method. One may see the shrinkage mechanism as a model selection device. To illustrate the point, we plot the number of estimated individual effects as a function of the tuning parameter in Figure 4.1 (panels (a) and (d)). The first panel illustrates how the degree of shrinkage (represented by the number of non-zero individual effects) changes with λ in our BHPS sample of male workers. For $\lambda \to 0$, $\hat{\alpha}_i(\lambda) \approx \hat{\alpha}_i(0) \approx \hat{\alpha}_i$, and the estimated conditional quantile function is,

$$\hat{Q}_{W_{it}}(\tau|\boldsymbol{x}_{it}, \alpha_i) = \boldsymbol{x}'_{it}\hat{\boldsymbol{\beta}}(\tau, \lambda) + \hat{\alpha}_i(\lambda) \approx \boldsymbol{x}'_{it}\hat{\boldsymbol{\beta}}(\tau, 0) + \hat{\alpha}_i(0),$$

which represents an estimated quantile regression model with individual effects. On the other hand, for $\lambda \approx 12$, we have that $\hat{\alpha}_i(\lambda) = 0$ for all male workers in our BHPS sample, implying that the estimated quantile regression model is,

$$\hat{Q}_{W_{it}}(\tau|\boldsymbol{x}_{it}, \alpha_i) = \boldsymbol{x}'_{it}\hat{\boldsymbol{\beta}}(\tau, \lambda) + \hat{\alpha}_i(\lambda) \approx \boldsymbol{x}'_{it}\hat{\boldsymbol{\beta}}(\tau).$$

This model corresponds to employing classical quantile regression techniques (see, e.g., Koenker 2005) on a panel data model. Therefore, panel (a) in Figure 4.1 reveals the number of individual effects α_i that are estimated from $\lambda \to 0$ (akin to fixed effects) through $\lambda \approx 12$ (akin to pooled cross-section). This range of values of λ's and the associated $\boldsymbol{\beta}$'s raises the following question: What is the optimal value of λ? We find that the modified AIC approach and the optimal shrinkage method described in Section 2.1 agree that λ should be relatively small. The values of $\hat{\lambda} = \{0.5, 0.2\}$ shown in the bottom panel of Table 4.1 are obtained by the optimal shrinkage method described in equation 2.7 and the modified AIC method.[9]

[8] Strictly speaking, the penalized least squares estimator is the estimator most closely associated with the penalized quantile regression approach. We refer the reader to Ruppert, Wand, and Carroll (2003) for details on penalized least squares methods for linear and non-linear models.

[9] As expected in panels (a) and (d) of Figure 4.1, we observe discontinuities in the number of non-zero individual effects. A small increment in the value of the tuning parameter λ shrinks the $\hat{\alpha}_i$'s toward zero, but the resulting $\hat{\alpha}_i$'s may not be exactly zero. For moderate increases in λ, we

	$\hat{\lambda}$	Quantiles 0.10	0.25	0.50	0.75	0.90	Mean
				Pooled Methods			
Intermediate qualifications	12	0.158*	0.177*	0.209*	0.235*	0.225*	0.207*
		(0.028)	(0.022)	(0.021)	(0.026)	(0.028)	(0.040)
Advanced qualifications	12	0.566*	0.617*	0.578*	0.564*	0.486*	0.556*
		(0.043)	(0.033)	(0.032)	(0.039)	(0.043)	(0.059)
Experience	12	0.040*	0.034*	0.037*	0.034*	0.034*	0.038*
		(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.006)
Experience Squared/100	12	-0.099*	-0.083*	-0.086*	-0.073*	-0.077*	-0.091*
		(0.009)	(0.006)	(0.006)	(0.007)	(0.007)	(0.012)
Married	12	0.172*	0.169*	0.148*	0.127*	0.145*	0.153*
		(0.018)	(0.014)	(0.013)	(0.016)	(0.018)	(0.027)
Non-English Speaking	12	-0.136*	-0.242*	-0.170*	-0.105†	-0.169*	-0.181†
Foreign-Born		(0.047)	(0.036)	(0.035)	(0.043)	(0.046)	(0.079)
English Speaking	12	0.112‡	0.243*	0.294*	0.328*	0.478*	0.305†
Foreign-Born		(0.065)	(0.050)	(0.049)	(0.060)	(0.064)	(0.139)
Location Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Individual Effects		No	No	No	No	No	No
Observations		8366	8366	8366	8366	8366	8366
				Panel Data Methods			
Intermediate qualifications	0.5	0.225*	0.222*	0.209*	0.202*	0.197*	0.221*
	[0.2]	(0.052)	(0.048)	(0.048)	(0.047)	(0.055)	(0.040)
Advanced qualifications	0.5	0.628*	0.639*	0.610*	0.600*	0.561*	0.606*
	[0.2]	(0.087)	(0.065)	(0.063)	(0.061)	(0.068)	(0.071)
Experience	0.5	0.057*	0.047*	0.040*	0.037*	0.038*	0.052*
	[0.2]	(0.007)	(0.004)	(0.004)	(0.005)	(0.006)	(0.003)
Experience Squared/100	0.5	-0.127*	-0.102*	-0.088*	-0.083*	-0.086*	-0.115*
	[0.2]	(0.017)	(0.010)	(0.010)	(0.011)	(0.012)	(0.007)
Married	0.5	0.076*	0.050*	0.041*	0.040†	0.048†	0.055*
	[0.2]	(0.024)	(0.018)	(0.018)	(0.019)	(0.021)	(0.013)
Non-English Speaking	0.5	-0.191‡	-0.177	-0.149	-0.144	-0.110	-0.137‡
Foreign-Born	[0.2]	(0.112)	(0.111)	(0.102)	(0.103)	(0.106)	(0.083)
English Speaking	0.5	0.395*	0.429*	0.423*	0.406*	0.384*	0.331*
Foreign-Born	[0.2]	(0.119)	(0.117)	(0.116)	(0.112)	(0.113)	(0.098)
Location Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Individual Effects		Yes	Yes	Yes	Yes	Yes	Yes
Observations		8366	8366	8366	8366	8366	8366

TABLE 4.1. *Comparison of pooled and panel data results from a BHPS sample of male workers. Mean refers to OLS and random effects estimators. The model also includes an intercept, age, and a quadratic term on age. The symbols ‡,†,∗ denote statistically different from zero at the 0.10, 0.05, and 0.01 level of significance.*

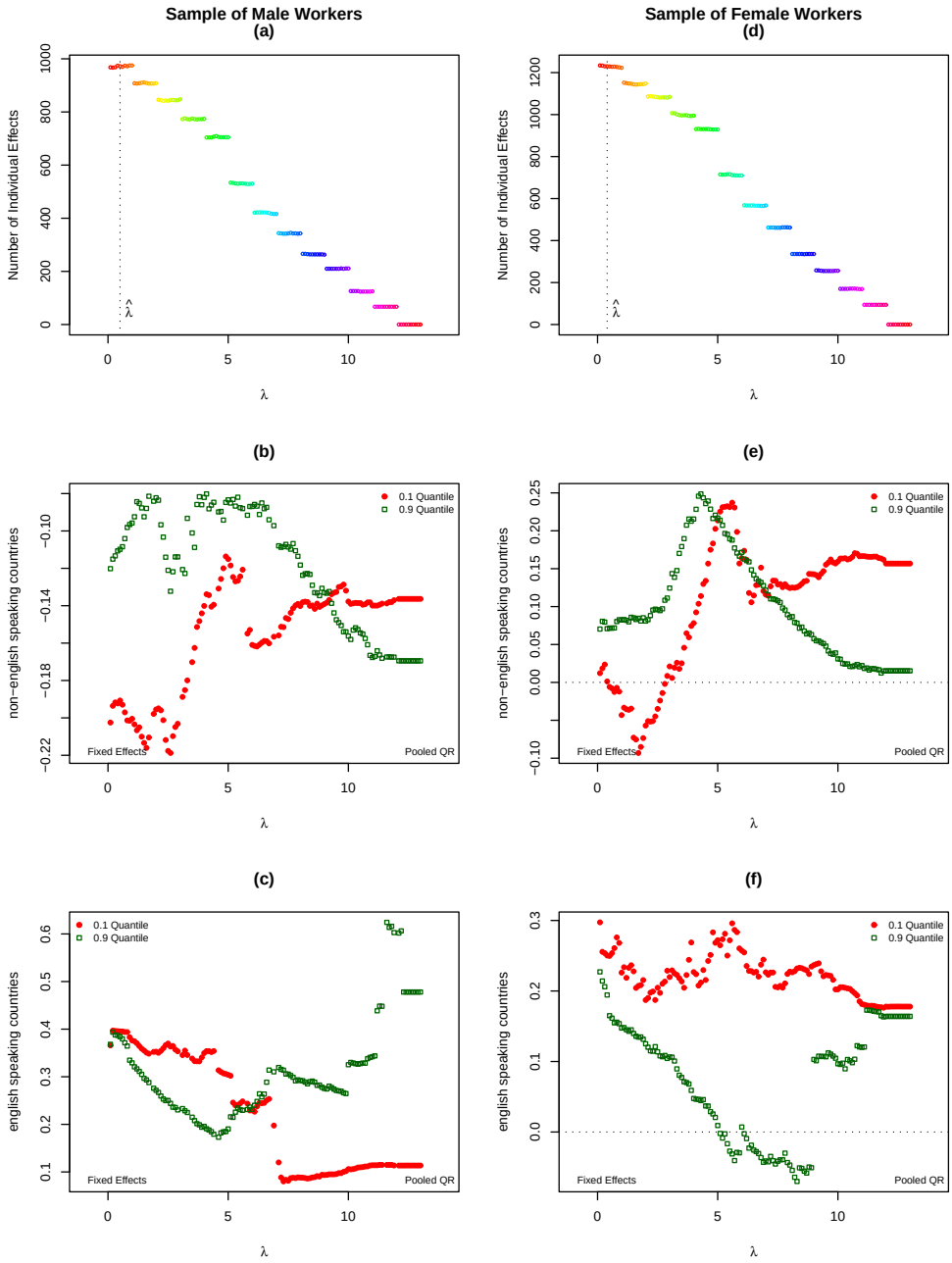

FIGURE 4.1. *Understanding the effect of immigrant unobserved heterogeneity in the BHPS sample. While the top panels show the number of non-zero individual effects as a function of* λ*, the other panels show the range of the estimated earnings differentials between immigrant and natives. For plotting these ranges, we consider the lower and upper tails of the conditional earnings distributions. Fixed effects estimates are shown at* $\lambda \approx 0$ *and pooled quantile regression estimates at* $\lambda = 12$*. The optimal shrinkage parameter estimate is indicated by* $\hat{\lambda}$*.*

It is interesting to see how the effects of being born in a non-English speaking country and being born in an English speaking country vary according to the shrinkage parameter λ. Panels (b) and (c) in Figure 4.1 present the estimated effects at the lower and upper tails of the conditional earnings distribution, obtained from the model estimated in Table 4.1. Panel (b) suggests that the heterogeneous earnings differential between immigrants born in non-English speaking countries and natives is not associated with language skills. After we control for unobserved heterogeneity, we find that the differentials tend to persists even with $\lambda \to 0$. However, the results of panel (c) indicate that the differentials across the conditional earnings distribution may be associated with unobserved language skills. The difference between the estimated effect at the 0.1 quantile and the estimated effect at the 0.9 quantile of the conditional earnings distribution disappears when the $\hat{\alpha}_i$'s are non-zero for most of the workers. This reveals the value of combining quantile regression with controls for unobserved heterogeneity.

Returning to the Table 4.1, we first examine the results for male workers in the BHPS in the upper block. We see that the mean results are relatively informative on the effects of the intermediate and advanced qualifications on earnings. For instance, while the mean difference between earnings of workers with intermediate qualifications and no qualifications is $\exp(0.207) - 1 \approx 0.230$ or 23.0 percent, the difference at the 0.1 quantile is 17.1 percent. However, the real advantage of our approach in this context can be seen when we investigate the effect of being born in an English speaking country. The mean effect suggests a 35.7 percent earnings differential between workers born in an English speaking country and natives, while the effects at the 0.1 quantile and 0.9 quantile indicate significantly different returns of 11.9 percent and 61.3 percent. It is interesting to see that the earnings differentials change sign by country of origin. The effect of being born in a non-English speaking country is negative and significant across the quantiles of the conditional distribution of earnings, but the estimated effect of being born in an English-speaking country is positive and significant across the quantiles. On the other hand, we are cautious in

expect the smallest individual effects to be zero. The behavior of the estimates plotted in the top panels is consistent with the evidence presented in Koenker (2004).

	$\hat{\lambda}$	Quantiles 0.10	0.25	0.50	0.75	0.90	Mean
				Pooled Methods			
Intermediate qualifications	12	0.488*	0.419*	0.320*	0.354*	0.470*	0.404*
		(0.050)	(0.040)	(0.024)	(0.023)	(0.020)	(0.056)
Advanced qualifications	12	1.395*	0.984*	0.890*	0.892*	0.919*	1.067*
		(0.108)	(0.090)	(0.055)	(0.053)	(0.046)	(0.108)
Experience	12	0.000	-0.023	-0.008†	0.010	0.029*	-0.003
		(0.008)	(0.007)	(0.004)	(0.004)	(0.003)	(0.007)
Experience Squared/100	12	-0.015	0.025	-0.003	-0.041*	-0.079*	-0.011
		(0.017)	(0.014)	(0.008)	(0.008)	(0.007)	(0.016)
Married	12	-0.290*	-0.332*	-0.203*	-0.107*	-0.112*	-0.193*
		(0.038)	(0.030)	(0.018)	(0.017)	(0.015)	(0.037)
Non-English Speaking	12	0.136	0.114	0.030	0.060	0.021	0.095
Foreign-Born		(0.085)	(0.071)	(0.043)	(0.041)	(0.036)	(0.082)
English Speaking	12	0.195	0.424*	0.146‡	0.020	0.165*	0.209
Foreign-Born		(0.137)	(0.114)	(0.069)	(0.066)	(0.056)	(0.151)
Location Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Individual Effects		No	No	No	No	No	No
Observations		10373	10373	10373	10373	10373	10373
				Panel Data Methods			
Intermediate qualifications	0.4	0.351*	0.313*	0.275*	0.247*	0.214*	0.406*
	[0.6]	(0.093)	(0.085)	(0.078)	(0.075)	(0.073)	(0.059)
Advanced qualifications	0.4	1.176*	0.988*	0.925*	0.883*	0.826*	1.111*
	[0.6]	(0.180)	(0.177)	(0.168)	(0.166)	(0.159)	(0.110)
Experience	0.4	0.008	0.006	0.000	-0.002	-0.004	0.001
	[0.6]	(0.007)	(0.005)	(0.005)	(0.005)	(0.007)	(0.004)
Experience Squared/100	0.4	-0.019	-0.022*	-0.019*	-0.019*	-0.017	-0.018*
	[0.6]	(0.013)	(0.008)	(0.009)	(0.009)	(0.013)	(0.007)
Married	0.4	-0.179*	-0.129*	-0.084*	-0.066*	-0.048†	-0.135*
	[0.6]	(0.031)	(0.019)	(0.015)	(0.015)	(0.020)	(0.018)
Non-English Speaking	0.4	0.001	0.080	0.081	0.091	0.071	0.150
Foreign-Born	[0.6]	(0.163)	(0.122)	(0.117)	(0.114)	(0.108)	((0.101)
English Speaking	0.4	0.250‡	0.227‡	0.196‡	0.186‡	0.194‡	0.238†
Foreign-Born	[0.6]	(0.137)	(0.113)	(0.111)	(0.108)	(0.098)	(0.135)
Location Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Individual Effects		Yes	Yes	Yes	Yes	Yes	Yes
Observations		10373	10373	10373	10373	10373	10373

TABLE 4.2. *Comparison of pooled and panel data results from a BHPS sample of female workers. Mean refers to OLS and random effects estimators. The model also includes an intercept, age, and a quadratic term on age. The symbols ‡,†,∗ denote statistically different from zero at the 0.10, 0.05, and 0.01 level of significance.*

interpreting this result, because unobserved language ability could be correlated with these independent variables.

We therefore introduce individual effects and reestimate the models. The results are presented in the lower panel of Table 4.1. We continue to see negative signs for the effect of being born in a non-English speaking country, and positive signs of the effect of being born in an English-speaking country. We note, however, that some of the effects are not significant. The results suggest that the effect of language on earnings is important at the lower tail of the conditional distribution of earnings. While being born in a non-English speaking country is associated with a reduction of 17.4 percent in earnings, being born in an English speaking country is associated with an increase in 48.4 percent in earnings. Moreover, it is interesting to see that while the effect of being born in an English speaking country tends to increase across quantiles in the model without individual effects, it tends to be constant in the model with individual effects.

Table 4.2 presents estimates for female workers. The results from pooled methods indicate no differences between immigrants born in non-English speaking countries and native-born workers. However, there seem to be significant positive differences between immigrants from English speaking countries and natives at the 0.25, 0.5 and 0.9 quantiles of the conditional distribution of earnings. These results are robust to the inclusion of individual effects in the model. The effect of being born in English speaking countries on earnings ranges from a positive and significant 28.4 at the 0.1 quantile to a significant 21.4 percent at the 0.9 quantile.

Turning to the effects of the other independent variables, it is interesting to observe that the effects of intermediate and advanced qualifications are reduced at the upper tail when we introduce individual effects. The earnings differential attributed to intermediate qualifications relative to no qualifications is 60.0 percent at the 0.9 quantile, which is dramatically reduced to 23.9 percent in the model estimated in the bottom part of Table 4.2.

4.1.1. *Additional empirical evidence.* Table 4.3 offers panel data estimates from models that incorporate additional covariates that are typically considered in the literature. The upper block presents results for the sample of male workers, and the bottom offers results associated with the sample of female workers. In addition to the

	Quantiles					
	0.10	0.25	0.50	0.75	0.90	Mean
			Male Sample			
Intermediate qualifications	0.232*	0.230*	0.207*	0.193*	0.191*	0.239*
	(0.054)	(0.051)	(0.051)	(0.052)	(0.056)	(0.041)
Advanced qualifications	0.642*	0.643*	0.611*	0.590*	0.568*	0.622*
	(0.081)	(0.067)	(0.064)	(0.064)	(0.070)	(0.072)
Experience	0.054*	0.046*	0.038*	0.037*	0.037*	0.049*
	(0.007)	(0.005)	(0.005)	(0.006)	(0.006)	(0.003)
Experience Squared/100	-0.118*	-0.098*	-0.083*	-0.081*	-0.082*	-0.107*
	(0.017)	(0.012)	(0.012)	(0.013)	(0.013)	(0.007)
Married	0.068*	0.045*	0.035‡	0.038‡	0.038‡	0.047*
	(0.020)	(0.018)	(0.020)	(0.020)	(0.022)	(0.013)
Non-English Speaking	-0.698‡	-0.727‡	-0.503	-0.432	-0.471	-0.367
Foreign-Born	(0.377)	(0.384)	(0.326)	(0.315)	(0.313)	(0.231)
English Speaking	-0.250	-0.199	0.024	0.073	-0.039	-0.079
Foreign-Born	(0.383)	(0.401)	(0.346)	(0.37)	(0.323)	(0.252)
Hours	0.009*	0.007*	0.006†	0.006†	0.006†	0.012*
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.001)
Number of children	0.004	0.009	0.009	0.006	0.011	0.007
	(0.009)	(0.007)	(0.006)	(0.007)	(0.008)	(0.006)
Union	0.055*	0.028*	0.011	0.005	-0.012	0.035*
	(0.019)	(0.012)	(0.014)	(0.018)	(0.024)	(0.010)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Location dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8366	8366	8366	8366	8366	8366
			Female Sample			
Intermediate qualifications	0.203*	0.254*	0.252*	0.255*	0.243*	0.290*
	(0.056)	(0.049)	(0.049)	(0.049)	(0.051)	(0.033)
Advanced qualifications	0.704*	0.737*	0.764*	0.796*	0.803*	0.791*
	(0.104)	(0.094)	(0.096)	(0.096)	(0.099)	(0.076)
Experience	0.026*	0.023*	0.020*	0.021*	0.019*	0.025*
	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.003)
Experience Squared/100	-0.062*	-0.059*	-0.055*	-0.058*	-0.056*	-0.067*
	(0.009)	(0.007)	(0.007)	(0.007)	(0.010)	(0.006)
Married	0.002	-0.009	-0.007	-0.014	-0.020	-0.002
	(0.019)	(0.014)	(0.013)	(0.014)	(0.017)	(0.013)
Non-English Speaking	-0.037	0.011	0.087	0.102	0.156	0.051
Foreign-Born	(0.328)	(0.232)	(0.224)	(0.219)	(0.229)	(0.169)
English Speaking	0.144	0.098	0.126	0.140	0.150	0.122
Foreign-Born	(0.260)	(0.163)	(0.160)	(0.155)	(0.169)	(0.133)
Hours	0.040*	0.039*	0.037*	0.035*	0.032*	0.038*
	(0.002)	(0.001)	(0.001)	(0.002)	(0.003)	(0.001)
Number of children	-0.081*	-0.063*	-0.052*	-0.048*	-0.040*	-0.073*
	(0.014)	(0.010)	(0.011)	(0.012)	(0.016)	(0.007)
Union	0.090*	0.047*	0.027	0.016	0.005	0.071*
	(0.018)	(0.016)	(0.022)	(0.026)	(0.030)	(0.011)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Location Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10373	10373	10373	10373	10373	10373

TABLE 4.3. *Panel data results from BHPS samples. The symbols ‡,†,∗ denote statistically different from zero at the 0.10, 0.05, and 0.01 level of significance. Other controls include years since migration, and race.*

independent variables used before, we introduce years since migration to the U.K., hours worked, number of children, and union membership. Additionally, the models include controls for workers race classified as Black, Indian, Pakistani or Bangladeshi, and Chinese and other groups. The indicator for whether worker's race is 'white' is omitted in the regressions.[10]

Our primary focus continues to be the native-born/immigrant earnings differentials. The results in the upper block of Table 4.3 suggest that male immigrants from non-English speaking countries are heavily penalized in terms of earnings, while immigrants from English speaking countries do not seem to experience any significant wage differential. This penalty tends to decrease in magnitude across higher quantiles of the conditional earnings distribution. On the other hand, the bottom panel of Table 4.3 suggests that there is no native-born/immigrant differential on earnings among female workers.

Further examining the results for men, we see that additional controls do not substantially alter the returns to education or experience, but in the full model we realize that the prior positive effect of marriage reflects a bias due to omitted variables. On the other hand, additional variables yield lower returns to schooling for women at the lower quantiles. Experience generates significant benefits and the marital penalty disappears. Increasing hours worked benefits women far more than men, as does union status. Perhaps unsurprisingly, children have no effect on men's earnings, but yield lower wages for women.

4.2. **PSID sample.** We now present results from the PSID sample. Table 4.4 lists estimates for male workers and Table 4.5 presents results for female workers. These tables are similar to the tables presented for the BHPS sample. The upper panel

[10]It is standard in the literature to estimate models that account for immigrant arrival cohort (see, e.g., Lubotsky 2007). As a robustness check, we estimated a similar model to the one presented in Table 4.3, adding indicators for time of the arrival to the U.K. These variables were generated by grouping immigrants into several groups. We consider immigrants who arrived between 1946 and 1962, immigrants who arrived between 1963 and 1972, and immigrants who arrived between 1973 and 1989. We obtained similar results. There is no native-born/immigrant differential on earnings among female workers. Moreover, while male workers who migrated from non-English speaking countries earn lower wages, workers who migrated from English speaking countries do not earn significantly different wages than native workers.

shows results from models that include the variables of interest, location dummies, and year dummies. The lower panel of the table displays results for models with the same covariates, location dummies and year dummies, as well as individual effects. In the last column, we present conditional mean results that are most closely associated with the quantile regression results.

Before turning to the empirical results, it is important to discuss the selection of the tuning parameter λ in Tables 4.4 and 4.5. As before, we plot the number of individual effects that are not equal to zero in the top panels of Figure 4.2. Notice that for $\lambda \approx 5.5$, we have $\hat{\alpha}_i(\lambda) \approx 0$ for all male and female workers. These results correspond to classical quantile regression techniques, because the method ignores individual heterogeneity. On the other hand, for $\lambda \to 0$, $\hat{\alpha}_i(\lambda) \approx \hat{\alpha}_i(0) \approx \hat{\alpha}_i$, representing a quantile regression version of the fixed effects model. The optimal value of the shrinkage parameter (determined as in equation 2.7) is 0.5 in Tables 4.4 and 4.5, similar to the value obtained by the AIC-type approach.

It is interesting to examine the sensitivity of the results to the choice of the tuning parameter λ. We briefly investigate the effect of the tuning parameter λ on the estimates corresponding to the coefficient associated with non-English speaking countries in our sample of male workers. By increasing λ, we can examine the results corresponding to different ways of addressing individual heterogeneity: from a fixed effects approach for $\lambda \approx 0$, to a classical quantile regression approach for $\lambda \geq 5.5$. For simplicity, we present in panel (b) of Figure 4.2 the effects corresponding to the $\{0.1, 0.9\}$ quantiles of the conditional distribution of earnings. The pooled quantile regression results indicate heterogeneous effects across the quantiles of the conditional earnings distribution, although this heterogeneity seems be attributed to unobserved differences in language ability and unobserved skills. The estimated differentials are much smaller when the tuning parameter λ tends to zero. The figure shows that some large native-born/immigrant differentials across quantiles suggested by the pooled approach are likely attributable to unobserved differences in language skills and ability.

In the PSID, the immigrant/native earnings gap is negative and significant (Table 4.4). For instance, workers on non-English speaking countries have earnings 43.3 percent lower than the native worker at the 0.1 quantile, and 31.4 percent lower than natives at the 0.9 quantile. Additionally, workers born in English speaking

	$\hat{\lambda}$	Quantiles 0.10	0.25	0.50	0.75	0.90	Mean
				Pooled Methods			
Intermediate qualifications	5.5	0.329*	0.350*	0.358*	0.424*	0.531*	0.396*
		(0.038)	(0.026)	(0.022)	(0.023)	(0.036)	(0.035)
Advanced qualifications	5.5	0.716*	0.801*	0.824*	0.881*	1.016*	0.851*
		(0.048)	(0.032)	(0.028)	(0.029)	(0.046)	(0.049)
Experience	5.5	0.027*	0.030*	0.036*	0.038*	0.033*	0.038*
		(0.005)	(0.004)	(0.003)	(0.003)	(0.005)	(0.005)
Experience Squared/100	5.5	-0.049*	-0.053*	-0.063*	-0.063*	-0.058*	-0.069*
		(0.011)	(0.008)	(0.007)	(0.008)	(0.011)	(0.010)
Married	5.5	0.256*	0.226*	0.202*	0.205*	0.215*	0.224*
		(0.029)	(0.020)	(0.017)	(0.017)	(0.028)	(0.028)
Non-English Speaking	5.5	-0.568*	-0.461*	-0.404*	-0.366*	-0.377*	-0.418*
Foreign-Born		(0.048)	(0.032)	(0.028)	(0.029)	(0.046)	(0.050)
English Speaking	5.5	-0.257‡	-0.422*	-0.423*	-0.413*	-0.450*	-0.370*
Foreign-Born		(0.146)	(0.105)	(0.091)	(0.093)	(0.145)	(0.116)
Location Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Individual Effects		No	No	No	No	No	No
Observations		10090	10090	10090	10090	10090	10090
				Panel Data Methods			
Intermediate qualifications	0.5	0.412*	0.380*	0.385*	0.385*	0.375*	0.402*
	[0.2]	(0.067)	(0.054)	(0.052)	(0.053)	(0.061)	(0.039)
Advanced qualifications	0.5	0.840*	0.816*	0.825*	0.838*	0.868*	0.858*
	[0.2]	(0.082)	(0.068)	(0.065)	(0.067)	(0.083)	(0.052)
Experience	0.5	0.064*	0.049*	0.041*	0.035*	0.031*	0.050*
	[0.2]	(0.011)	(0.008)	(0.005)	(0.007)	(0.010)	(0.004)
Experience Squared/100	0.5	-0.118*	-0.088*	-0.074*	-0.062*	-0.054*	-0.091*
	[0.2]	(0.023)	(0.015)	(0.010)	(0.013)	(0.019)	(0.008)
Married	0.5	0.113*	0.076†	0.056†	0.055‡	0.022	0.102*
	[0.2]	(0.053)	(0.034)	(0.024)	(0.032)	(0.049)	(0.021)
Non-English Speaking	0.5	-0.368*	-0.381*	-0.359*	-0.323*	-0.327*	-0.394*
Foreign-Born	[0.2]	(0.082)	(0.068)	(0.066)	(0.066)	(0.074)	(0.046)
English Speaking	0.5	-0.469	-0.491†	-0.445‡	-0.451†	-0.341	-0.382*
Foreign-Born	[0.2]	(0.315)	(0.211)	(0.210)	(0.233)	(0.315)	(0.126)
Location Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Individual Effects		Yes	Yes	Yes	Yes	Yes	Yes
Observations		10090	10090	10090	10090	10090	10090

TABLE 4.4. *Comparison of pooled and panel data results from a PSID sample of male workers. Mean refers to OLS and random effects estimators. The model also includes an intercept, age, and a quadratic term on age. The symbols ‡,†,∗ denote statistically different from zero at the 0.10, 0.05, and 0.01 level of significance.*

	$\hat{\lambda}$	Quantiles 0.10	0.25	0.50	0.75	0.90	Mean
				Pooled Methods			
Intermediate qualifications	5.5	0.396*	0.339*	0.356*	0.368*	0.388*	0.376*
		(0.047)	(0.035)	(0.031)	(0.031)	(0.048)	(0.047)
Advanced qualifications	5.5	0.960*	0.924*	0.932*	0.890*	0.828*	0.892*
		(0.057)	(0.043)	(0.038)	(0.038)	(0.059)	(0.058)
Experience	5.5	0.009	0.013*	0.015*	0.009†	0.015†	0.011*
		(0.006)	(0.005)	(0.004)	(0.004)	(0.006)	(0.005)
Experience Squared/100	5.5	-0.015	-0.025*	-0.029*	-0.019†	-0.033*	-0.022*
		(0.012)	(0.009)	(0.008)	(0.008)	(0.012)	(0.011)
Married	5.5	0.064†	0.097*	0.068*	0.067*	0.073*	0.064*
		(0.026)	(0.020)	(0.018)	(0.017)	(0.026)	(0.024)
Non-English Speaking	5.5	-0.296*	-0.359*	-0.315*	-0.297*	-0.262*	-0.301*
Foreign-Born		(0.063)	(0.047)	(0.042)	(0.042)	(0.065)	(0.072)
English Speaking	5.5	0.123	0.011	-0.191†	-0.332*	0.020	-0.031
Foreign-Born		(0.343)	(0.097)	(0.087)	(0.086)	(0.133)	(0.149)
Location Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Individual Effects		No	No	No	No	No	No
Observations		8790	8790	8790	8790	8790	8790
				Panel Data Methods			
Intermediate qualifications	0.5	0.452*	0.423*	0.360*	0.333*	0.305*	0.381*
	[0.6]	(0.087)	(0.072)	(0.065)	(0.064)	(0.076)	(0.056)
Advanced qualifications	0.5	1.050*	1.043*	0.991*	0.963*	0.941*	0.911*
	[0.6]	(0.106)	(0.084)	(0.079)	(0.078)	(0.091)	(0.068)
Experience	0.5	0.037*	0.032*	0.028*	0.019*	0.011	0.019*
	[0.6]	(0.011)	(0.008)	(0.006)	(0.008)	(0.010)	(0.004)
Experience Squared/100	0.5	-0.065*	-0.056*	-0.050*	-0.035†	-0.021	-0.039*
	[0.6]	(0.022)	(0.016)	(0.019)	(0.015)	(0.020)	(0.008)
Married	0.5	0.038	0.040	0.026	0.021*	0.032	0.037‡
	[0.6]	(0.041)	(0.034)	(0.028)	(0.030)	(0.039)	(0.021)
Non-English Speaking	0.5	-0.395*	-0.396*	-0.359*	-0.363	-0.334*	-0.274*
Foreign-Born	[0.6]	(0.103)	(0.090)	(0.088)	(0.085)	(0.098)	(0.059)
English Speaking	0.5	-0.110	-0.179	-0.197	-0.179	-0.197	-0.005
Foreign-Born	[0.6]	(0.175)	(0.187)	(0.188)	(0.174)	(0.185)	(0.097)
Location Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Year Dummies		Yes	Yes	Yes	Yes	Yes	Yes
Individual Effects		Yes	Yes	Yes	Yes	Yes	Yes
Observations		8790	8790	8790	8790	8790	8790

TABLE 4.5. *Comparison of pooled and panel data results from a PSID sample of female workers. Mean refers to OLS and random effects estimators. The model also includes an intercept, age, and a quadratic term on age. The symbols ‡,†,∗ denote statistically different from zero at the 0.10, 0.05, and 0.01 level of significance.*

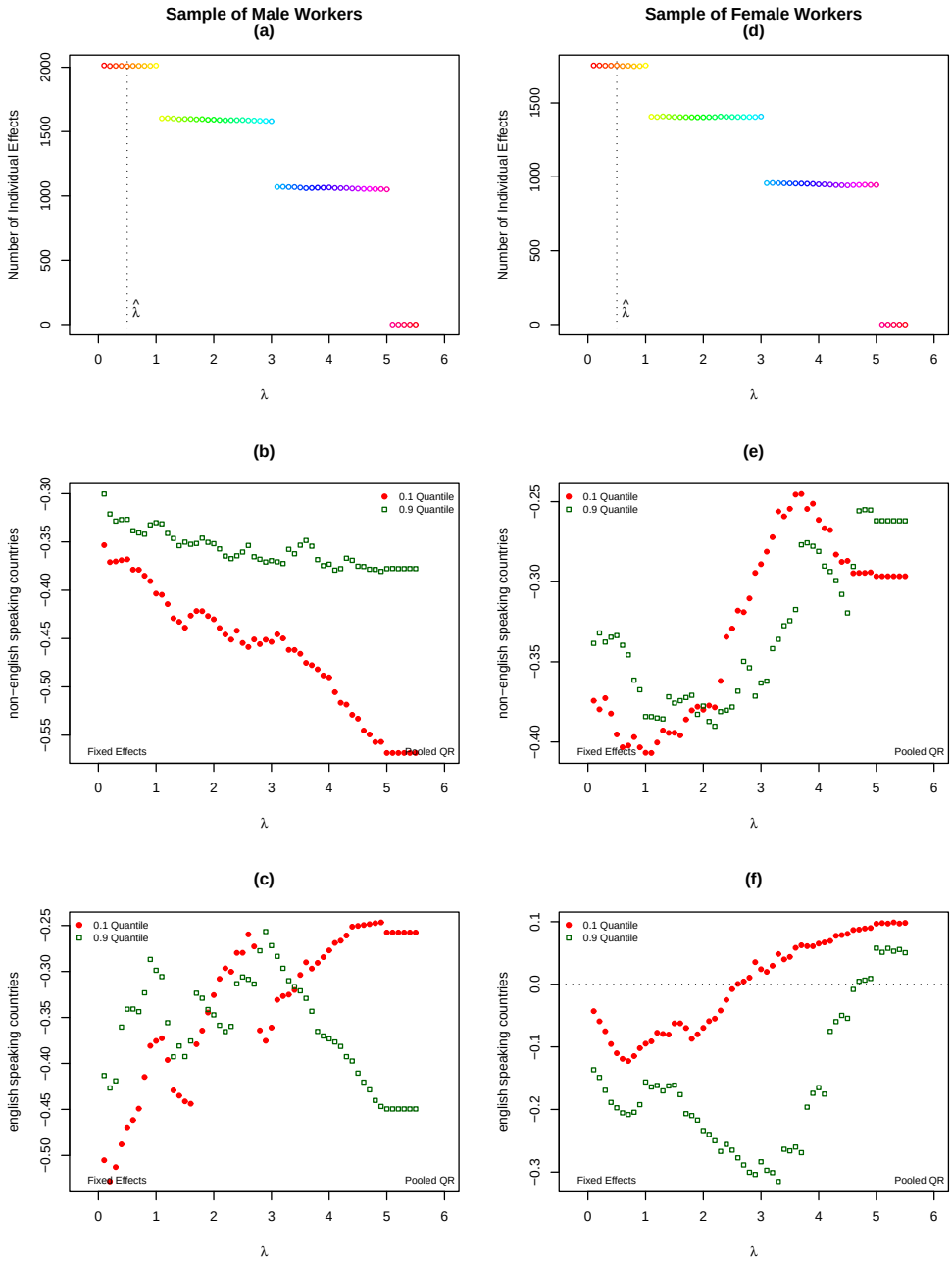

FIGURE 4.2. *Understanding the effect of immigrant unobserved heterogeneity in the PSID sample. While the top panels show the number of non-zero individual effects as a function of λ, the other panels show the range of the estimated earnings differentials between immigrant and natives. For plotting these ranges, we consider the lower and upper tails of the conditional earnings distributions. Fixed effects estimates are shown at $\lambda \approx 0$ and pooled quantile regression estimates at $\lambda = 5.5$. The optimal shrinkage parameter estimate is indicated by $\hat{\lambda}$.*

countries earn 22.7 percent less than natives at the 0.1 quantile, and 36.2 percent at the 0.9 quantile. These estimated earnings differentials continue to be negative and significant in the models with individual effects. (The sole exceptions are the effect of being born in an English speaking country at the 0.1 and 0.9 quantiles).

Table 4.4 shows larger earnings differentials attributed to educational qualifications than in the BHPS. The effects of intermediate qualifications are positive and significant across quantiles, ranging from 39.0 percent at the 0.1 quantile to 70.1 percent at the 0.9 quantile. The differences at the quantiles of the conditional distribution associated with advanced qualifications imply a 104.6 percent at the 0.1 quantile and 176.2 percent at the 0.9 quantile. When individual effects are included, the estimated effects at the tails are changed, suggesting a relatively constant return across quantiles of the earning distribution.

Finally, it is also interesting to see that the earnings differential associated with country of origin appears to be gender specific. Table 4.4 shows that immigrants earn less than natives, but Table 4.5 suggests that only the effect of emigrating from a non-English speaking country is significant for women. This result is robust to the inclusion of additional variables, as shown in the next section.

4.2.1. *Additional empirical evidence.* In Table 4.6, we offer panel data results from a model that includes additional controls. The upper panel of the table presents results from a sample of male workers, and the lower part lists estimates for female workers. As in the case of the British sample, we include years since migration, controls for workers race, hours worked, number of children and union membership. There is no significant wage penalty for male immigrants from English-speaking countries at the lower and upper tails, and the returns to emigrating from a non-English speaking country are negative and significant. (The sole exception is at the 0.1 quantile). The findings in the bottom panel of Table 4.6 suggest that native-immigrant earnings differentials in the lower tail for women are associated with immigrant status. The estimated effects associated with being born in a non-English country are negative and significant across most of the quantiles of the conditional distribution of earnings. The evidence suggests that immigrants who have weak English proficiency receive lower wages. The estimated effects associated with being born in an English speaking

	Quantiles					
	0.10	0.25	0.50	0.75	0.90	Mean
			Male Sample			
Intermediate qualifications	0.382*	0.358*	0.353*	0.359*	0.356*	0.353*
	(0.054)	(0.046)	(0.042)	(0.042)	(0.047)	(0.038)
Advanced qualifications	0.782*	0.779*	0.781*	0.792*	0.822*	0.779*
	(0.081)	(0.070)	(0.065)	(0.063)	(0.065)	(0.052)
Experience	0.065*	0.050*	0.042*	0.037*	0.035*	0.051*
	(0.007)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
Experience Squared/100	-0.118*	-0.087*	-0.071*	-0.065*	-0.061*	-0.093*
	(0.015)	(0.008)	(0.007)	(0.008)	(0.009)	(0.009)
Married	0.095*	0.042†	0.031‡	0.028‡	-0.005	0.079*
	(0.030)	(0.020)	(0.016)	(0.017)	(0.028)	(0.021)
Non-English Speaking	-0.306	-0.438*	-0.395*	-0.420*	-0.574*	-0.401*
Foreign-Born	(0.202)	(0.152)	(0.157)	(0.162)	(0.179)	(0.131)
English Speaking	-0.449	-0.628*	-0.574*	-0.642*	-0.699	-0.486*
Foreign-Born	(0.277)	(0.246)	(0.244)	(0.265)	(0.462)	(0.164)
Hours	-0.007*	-0.009*	-0.010*	-0.011*	-0.011*	-0.008*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of children	0.001	0.015*	0.019*	0.022*	0.027*	0.020*
	(0.011)	(0.006)	(0.005)	(0.005)	(0.008)	(0.007)
Union	0.155*	0.102*	0.073*	0.052*	0.011	0.104*
	(0.022)	(0.018)	(0.017)	(0.018)	(0.021)	(0.019)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Location dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10090	10090	10090	10090	10090	10090
			Female Sample			
Intermediate qualifications	0.388*	0.316*	0.252*	0.219*	0.230*	0.319*
	(0.064)	(0.060)	(0.049)	(0.047)	(0.046)	(0.056)
Advanced qualifications	0.885*	0.864*	0.802*	0.772*	0.787*	0.823*
	(0.093)	(0.079)	(0.068)	(0.065)	(0.064)	(0.068)
Experience	0.034*	0.025*	0.022*	0.014*	0.008	0.022*
	(0.009)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Experience Squared/100	-0.070*	-0.054*	-0.047*	-0.032*	-0.022†	-0.049*
	(0.016)	(0.010)	(0.010)	(0.009)	(0.011)	(0.009)
Married	0.039	0.024	0.028	0.021	0.006	0.004
	(0.039)	(0.031)	(0.026)	(0.024)	(0.026)	(0.022)
Non-English Speaking	-0.585†	-0.470*	-0.494*	-0.403‡	-0.297	-0.495*
Foreign-Born	(0.285)	(0.193)	(0.184)	(0.208)	(0.192)	(0.184)
English Speaking	-0.241	-0.254	-0.201	-0.163	-0.121	-0.144
Foreign-Born	(0.275)	(0.184)	(0.166)	(0.176)	(0.165)	(0.182)
Hours	-0.001	-0.002†	-0.003*	-0.006*	-0.008*	-0.008*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Number of children	-0.059*	-0.039*	-0.026*	-0.019*	-0.016	-0.026*
	(0.012)	(0.010)	(0.008)	(0.008)	(0.010)	(0.008)
Union	0.168*	0.177*	0.165*	0.120*	0.108*	0.092*
	(0.035)	(0.029)	(0.023)	(0.023)	(0.026)	(0.019)
Other controls	Yes	Yes	Yes	Yes	Yes	Yes
Location dummies	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Individual effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	8790	8790	8790	8790	8790	8790

TABLE 4.6. *Panel data results from PSID samples. The symbols ‡,†,∗ denote statistically different from zero at the 0.10, 0.05, and 0.01 level of significance. Other controls include years since migration, and race.*

country also suggest that immigrants earn lower wages, although these effects are not significant.

As with the British data, the full model reveals very similar returns to education and experience for men in the U.S. Immigrants with more years since immigration see higher earnings, particularly in the upper quantiles. High-earning immigrant men appear to have gained U.S.-specific skills that are valued in the labor market. The U.S. data show significant wage differences across racial and ethnic groups, as expected. Male workers in the U.S. earn higher wages when they have children and if they belong to a union.

Turning to women in the U.S., the returns to schooling are lower again in the full model. Here, marital status yields no penalty, but each child lowers a woman's wage at the lowest quantiles. The return to years since migration is also higher for women than men in the U.S. These results suggest circumstances by which the female immigration penalty is mitigated for some. Racial and ethnic wage differentials persist. We also see that female workers earn higher wages if they belong to a union, with the largest significant effect at the lower tail of the conditional earnings distribution.[11]

4.3. **Wage penalties across the conditional earnings distribution.** To summarize a few important findings, we present four graphs describing the profile of wage penalties by country of origin and gender. Figure 4.3 offers an opportunity to investigate in detail the effects of interest using panel data quantile regression. The panels on the left present evidence from the BHPS sample (Tables 4.1, 4.2 and 4.3), and the panels on the right present evidence from the PSID sample (Tables 4.4, 4.5, and 4.6). To improve the presentation, we do not include 95 percent confidence intervals, but the interested reader can find the standard errors in the previous tables.

[11]As in the case of the British sample, we estimated models that account for immigrant arrival cohort. Our model included indicators for time of arrival to the U.S., in addition to years since migration, controls for workers race, hours worked, number of children and union membership. In contrast to the British sample, we see significant cohort effects in the U.S among male workers. In results not reported here, the cohort penalty is lower for younger groups. Turning to women in the U.S., cohort effects are often negative, as they were for men, but the actual penalties are much smaller in magnitude.

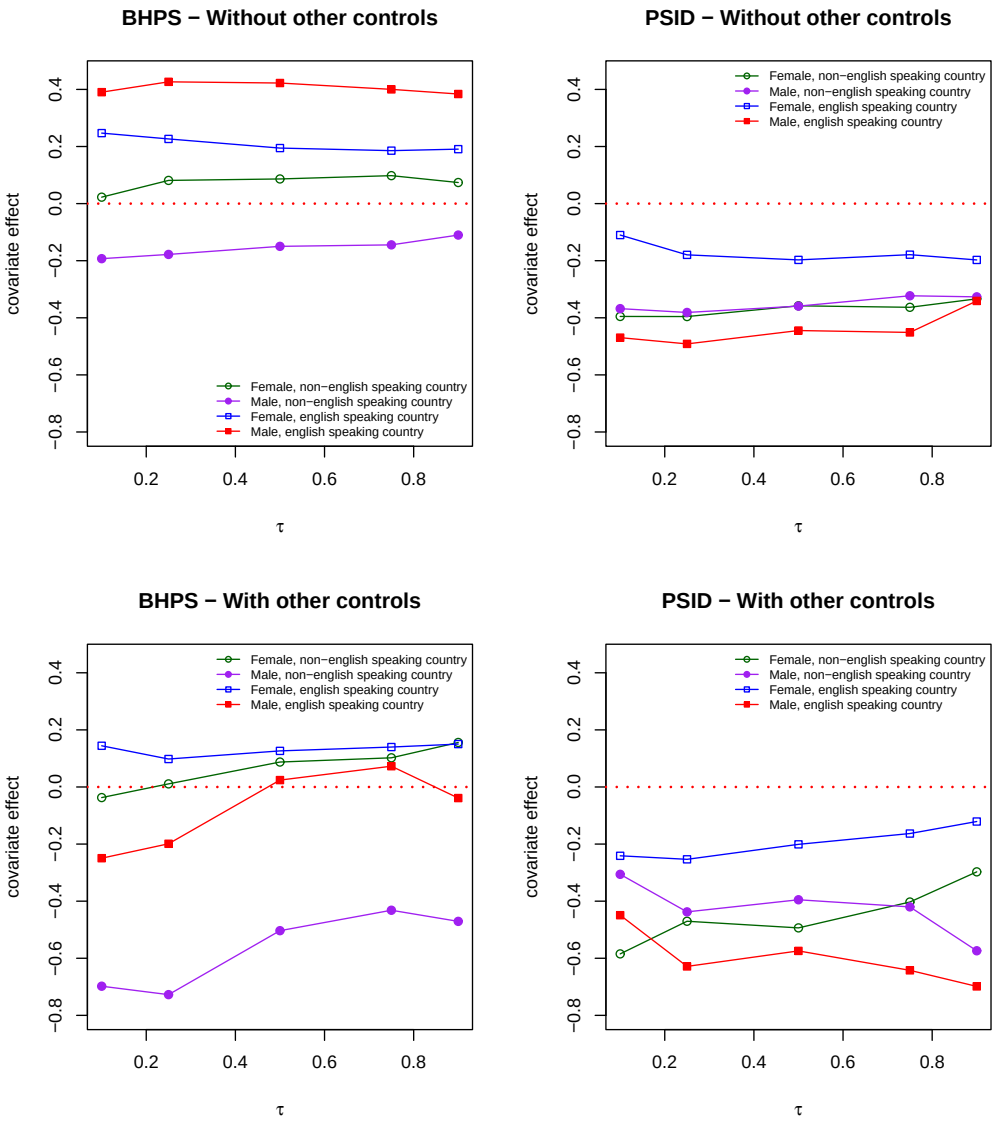

FIGURE 4.3. *An illustrative comparison of wage penalties by country of origin and gender in the United Kingdom and United States. The figure shows the profile of the covariate effects of interest at different quantiles τ of the conditional earnings distribution. These effects are estimated by penalized quantile regression. Other controls include years since migration, controls for workers race, hours worked, number of children and union membership.*

The graphs show that the wage penalty varies by country of origin and gender. In the BHPS sample, male workers from non-English speaking countries have lower earnings than natives, but male workers from English speaking countries do not have significantly lower earnings than natives. On the other hand, proficiency in English does not seem to be associated with higher wages in the PSID sample, because immigrant workers earn less than natives. Moreover, even after accounting for unobserved differences in language skills and ability among male workers, we observe that the effects of interest have a tendency to decrease through higher quantiles of the conditional earnings distribution. Finally, it is interesting to see that female workers who migrated to the U.S. seem to earn lower wages than native workers, while female workers who migrated to the U.K. do not experience a wage penalty relative to natives.

5. Conclusions

We use a relatively new panel data quantile regression method to explore immigrant-native wage differentials throughout the conditional earnings distribution, while controlling for individual heterogeneity. With data from the United Kingdom, we find that male immigrants from non-English-speaking countries often earn significantly lower wages than natives consistently throughout the (conditional) wage distribution. Immigrants who have weaker English proficiency receive substantially lower wages. Our evidence reveals that individual heterogeneity is important in determining earnings, and our panel data methods reveal robust earnings differentials at the mean and some additional points in the wage distribution. For instance, at the tenth decile of earnings, the impact of immigrating from a non-English speaking country is 17 percent lower earnings. This could highlight the importance of language skills among low-earners, that these immigrants select into lower-wage jobs, and/or the presence of wage discrimination. We also see that the parameters on immigrating from an English-speaking country become much less significant in our full model. The results for the female sample are quite different, revealing hardly any significant differential.

We also explore differentials among American immigrants and natives. In the U.S. immigrant men receive lower wages in nearly every case, regardless of country of

origin. It is particularly interesting to note the results from the panel data model; controlling for individual heterogeneity reveals that immigrants from English-speaking countries receive a larger wage penalty than those from countries like Mexico. Turning our attention to women in the U.S., we see that high-paid women from English countries receive earnings much like their native counterparts. Immigrant women from non-English countries earn much lower wages, substantially so in the panel data model. These negative differentials rise monotonically as wages rises, suggesting that perhaps these women face poor job prospects. Given the prevalence of women in the service sector, it is easy to believe that language ability is particularly important.

There are continuing debates about the effects of immigration and the role of immigrants in host countries. To aid our ongoing discovery about the work-life experiences of immigrants, we employ a new technique to more carefully understand earnings differentials. It is clear that wage differences remain, and vary a great deal among men and women, by country of origin, and across the earnings distribution.

References

Angrist, J., V. Chernozhukov, and I. Fernández-Val (2006): "Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure," *Econometrica*, 74(2), 539–563.

Bleakley, H., and A. Chin (2004): "Language Skills and Earnings: Evidence from Childhood Immigrants," *The Review of Economics and Statistics*, 86(2), 481–96.

Borjas, G. J. (1989): "Immigrant and Emigrant Earnings: A Longitudinal Study," *Economic Inquiry*, 27(1), 21–37.

Bratsberg, B., and J. F. Ragan (2002): "The Impact of Host-Country Schooling on Earnings: A Study of Male Immigrants in the United," *Journal of Human Resources*, 37(1), 63–105.

Butcher, K. F., and J. DiNardo (2002): "The Immigrant and Native-Born Wage Distributions: Evidence from United States Censuses," *Industrial and Labor Relations Review*, 56(1), 97–121.

Chiswick, B., A. T. Le, and P. W. Miller (2008): "How Immigrants Fare across the Earnings Distribution in Australia and the United States," *Industrial and Labor Relations Review*, 61(3), 353–73.

Chiswick, B., Y. L. Lee, and P. W. Miller (2005): "Immigrant Earnings: A Longitudinal Analysis," *Review of Income and Wealth*, 51(4), 485–503.

Chiswick, B., and P. W. Miller (1995): "The Endogeneity between Language and Earnings: International Analyses," *Journal of Labor Economics*, 13(2), 246–88.

——— (2007): "Earnings and Occupational Attainment: Immigrants and the Native Born," IZA Discussion Paper No. 2676.

——— (2009): "Earnings and Occupational Attainment among Immigrants," *Industrial Relations*, 48(3), 454–465.

Clark, K., and J. Lindley (2006): "Immigrant Labour Market Assimilation and Arrival Effects: Evidence from the UK Labour Force Survey," IZA Discussion Paper No. 2228.

Duleep, H. O., and D. J. Dowhan (2002): "Insights from Longitudinal Data on the Earnings Growth of U.S. Foreign-Born Men," *Demography*, 39(3), 458–506.

Dustmann, C., and A. V. Soest (2002): "Language and the Earnings of Immigrants," *Industrial and Labor Relations Review*, 55(3), 473–92.

Friedberg, R. M. (2000): "You Can't Take It with You? Immigrant Assimilation and the Portability of Human Capital," *Journal of Labor Economics*, 18(2), 221–51.

Galvao, A. (2009): "Quantile Regression for Dynamic Panel Data with Fixed Effects," mimeo, University of Illinois at Urbana-Champaign.

Graham, B. S., J. Hahn, and J. L. Powell (2009): "The Incidental Parameter Problem in a Non-differentiable Panel Data Model," *Economics Letters*, 105, 181–182.

Green, D. A. (1999): "Immigrant Occupational Attainment: Assimilation and Mobility over Time," *Journal of Labor Economics*, 17(1), 49–79.

Harding, M., and C. Lamarche (2009): "A Quantile Regression Approach for Estimating Panel Data Models Using Instrumental Variables," *Economics Letters*, 104, 133–135.

Hum, D., and W. Simpson (2004): "Reinterpreting the Performance of Immigrant Wages from Panel Data," *Empirical Economics*, 29, 129–47.

Kniesner, T. J., and J. D. Leeth (2010): "Hedonic Wage Equilibrium: Theory, Evidence and Policy," IZA Discussion Paper No. 5076.

Kniesner, T. J., W. K. Viscusi, and J. P. Ziliak (2010): "Policy Relevant Heterogeneity in the Value of Statistical Life: New Evidence from Panel Data Quantile Regressions," *Journal of Risk and Uncertainty*, 40(1), 15–31.

Koenker, R. (2004): "Quantile Regression for Longitudinal Data," *Journal of Multivariate Analysis*, 91, 74–89.

——— (2005): *Quantile Regression*. Cambridge University Press.

Lamarche, C. (2010): "Robust Penalized Quantile Regression Estimation for Panel Data," *Journal of Econometrics*, 157, 396–408.

Lubotsky, D. (2007): "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings," *Journal of Political Economy*, 115(5), 820–67.

Ruppert, D., M. P. Wand, and R. J. Carroll (2003): *Semiparametric Regression*. Cambridge University Press.

Weiss, Y., R. M. Sauer, and M. Gotlibovski (2003): "Immigration, Search, and Loss of Skill," *Journal of Labor Economics*, 21(3).