JOURNAL OF REGIONAL SCIENCE, VOL. 42, NO. 1, 2002, pp. 127–143

A METHODOLOGY FOR ESTIMATING RETURNS TO SKILLS FOR CANADIAN PROVINCES AND U.S. STATES*

Gary L. Hunt

Department of Economics, University of Maine, Orono, ME 04469, U.S.A. Email: gary.hunt@umit.maine.edu

Richard E. Mueller

Department of Economics, University of Lethbridge, Lethbridge, Alberta T1K 3M4, Canada. Email: richard.mueller@uleth.ca

ABSTRACT. Differences in both regional skill prices and skill mix can explain interregional variations in wage distributions. We control for interregional differences in skill mix that permit us to compute key parameters of regional wage distributions including regional returns to skills. In addition to setting forth the methods in detail, we also present estimates for 48 U.S. states and 10 Canadian provinces. For both males and females, we find that regional mean wages (with controls for skills mix) in the U.S. and Canada are similar, but that the returns to skills are systematically higher in the U.S.

1. INTRODUCTION

The purpose of this study is to describe a method for estimating key parameters of regional wage distributions, which include regional returns to skills, and to present estimates of these parameters for 48 U.S. states and 10 Canadian provinces in the early 1990s. Various studies have found variations in wages across regions in the United States and Canada (e.g., Hanushek, 1973; Dickie and Gerking, 1987, 1998). Such differences are attributable to interregional differences in skill composition, in returns to skills, or both. Studies pre-dating the spatial general equilibrium model of Roback (1982) often attributed interregional differences in measured returns to skills to disequilibrium forces or barriers to factor mobility. Spatial equilibrium studies have found evidence that nonconvergence in interregional wages also can be interpreted, in part, as an equilbrium phenomenon reflecting compensating differentials arising from heterogeneity in the spatial distribution of amenities (e.g., Roback, 1982, 1988; Blomquist, Berger, and Hoehn, 1988; Beeson and Eberts, 1989; Beeson, 1991; Greenwood et al., 1991).

The basic approach of most of the spatial equilibrium style studies is to estimate log wage equations on individual data and then to test for differential

* We thank the three anonymous referees for useful suggestions.
Received May 2000; revised January 2001; accepted March 2001.

Blackwell Publishing, Inc., 350 Main Street, Malden, MA 02148, USA and 108 Cowley Road, Oxford, OX4 1JF, UK.

regional intercepts or returns to skill parameters (e.g., coefficients on education). Results are then presented for these interregional pooling tests, and as indicated above, typically find that common regional intercepts or returns to skills can be rejected at conventional levels of statistical significance. In some studies (e.g., Beeson, 1991), additional variables are added to the log wage equations to test whether the significant interregional differences are accounted for by interregional variations in amenities, industrial structure, or demand conditions.

Our study uses this basic approach to estimate log wage equations for each of 58 North American regions using micro data on individual workers residing in 48 U.S. states and 10 Canadian provinces. Also by pooling the entire sample across all regions, we develop estimates of the distribution of skills across these 58 North American areas as a whole. The estimated skills are individual-specific and do not vary by region. Therefore, they can be used as the basis to form an interregionally standardized skill distribution that is invariant by region. This standardized distribution permits us, in conjunction with our estimates of each region's log wage equation, to estimate both a location parameter and a returns to skills parameter for each of the 58 North American regions. By design, our estimates of these parameters are not inclusive of interregional variations in returns that could be due to interregional variations in amenities or demand conditions. Specific estimates of such regional parameters have not been reported in the regional science literature previously, although several studies referenced above have formally tested, and found support for, interregional variations in returns to skill. Specific estimates of the geographic pattern of returns to skills are important in the study of several issues in regional science including interregional convergence and compensating differentials (see above), interregional migration (Borjas, Bronars, and Trejo, 1992; Hunt and Mueller, 2000), and human-capital based endogenous regional growth (Rauch, 1993). Consequently, both the method described here and the regional returns to skills parameter estimates reported here should be useful in several research programs in regional science and economics.

In the following section, we describe the theoretical and empirical aspects of the methodology in detail. In Section 3 we discuss the data necessary to implement our estimates. The econometric-based estimates are presented and discussed in Section 4. The final section concludes the paper.

2. METHODOLOGY

Theoretical Model

Following Borjas, Bronars, and Trejo (1992), we write the natural logarithm of an individual's wage in region j as

$$\ln(w_{ij}) = \mu_j + \phi_j(\upsilon_i - \upsilon) \tag{1}$$

where μ_j is the mean log wage in area j, ϕ_j is the return to skills in area j, υ_i is the individual's skill level, and υ is the global mean skill level (i.e., that for all individuals in all regions). Equation (1) assumes that the wage determination process can be characterized by a "one-factor" model of ability and that the position of each individual in the global skills distribution (i.e., the value of $\upsilon_i - \upsilon$) is interregionally invariant. In other words, Corr $(\upsilon_{ij}, \upsilon_{ik}) = 1, j \neq k$, where j and k index regions. Therefore, we require no skills index subscripts for area. Individual skills υ_i are independent of area.

Wage differences are generated by spatial variations in μ_j and ϕ_j. Taking the first two moments of Equation (1) with respect to individuals, we obtain

$$E[\ln(w_{ij})] = \mu_j + \phi_j[E(\upsilon_i) - \upsilon]$$

$$\mathrm{Var}[\ln(w_{ij})] = \phi_j^2 \mathrm{Var}(\upsilon_i)$$

If the individuals in area j have above (below) average skills, then $E(\upsilon_i) >(<) \upsilon$. In such cases, the mean of the log wage distribution will differ across areas due to both interregional differences in average skills, $E(\upsilon_i) - \upsilon$, and the values of μ_j and ϕ_j. Interregional differences in the area-specific variance of the log wage distribution will occur because of differences in ϕ_j and $\mathrm{Var}(\upsilon_i)$.

We are interested in regional variations in returns to skills, so we need to remove variations in the interregional log wage distribution parameters that result from differences in the skills mix. In other words, we attribute differences between regions to differential returns to skills and not differences in skills. This is achieved by using a standardized skill distribution with $E(\upsilon_i) = \upsilon$, and $\mathrm{Var}(\upsilon_i) = \sigma^2$.

For this standardized distribution, the first two moments of the log wage distribution are

$$\begin{aligned} E\left[\ln(w_{ij})^*\right] &= \mu_j + \phi_j\left[E(\upsilon_i) - \upsilon\right] \\ &= \mu_j + \phi_j[\upsilon - \upsilon] \\ &= \mu_j \end{aligned} \tag{2}$$

$$\begin{aligned} \mathrm{Var}\left[\ln(w_{ij})^*\right] &= \phi_j^2 \mathrm{Var}(\upsilon_i) \\ &= \phi_j^2 \sigma^2 \end{aligned} \tag{3}$$

Given these estimates of the first two moments of the standardized log wage distribution, the values of μ_j and ϕ_j are identified as $\mu_j = E[\ln(w_{ij})^*]$ and $\phi_j = \{\mathrm{Var}[\ln(w_{ij})^*]/\sigma^2\}^{1/2}$. Substitution of these values into Equation (1) implies that individual i's log wage in area j depends on the mean and variance of the standardized log wage distribution, the variance of the skills distribution, and the individual's algebraic difference from the mean skill level. If we denote the latter as the individual's skill differential, an individual with a positive skill differential will have a higher log wage in an area with a higher value of ϕ_j, ceteris paribus. In contrast, an individual with a negative skill differential will have a higher log wage in an area with a lower value of ϕ_j, ceteris paribus. We now turn our attention to the empirical construction of the two area-specific parameters, μ_j and ϕ_j.

Empirical Implementation

Area Mean Log Wage (μ_j). Equation (2) above indicates that μ_j is equal to the expected value of the standardized log wage distribution for area *j*. We compute an estimate of this expectation for each of the 58 areas as follows. First, we specify the following log wage equation for individuals *i*

$$\begin{aligned} \ln(w_i) = {} & \alpha_0 + \beta_1 ED_i + \beta_2 PX_i + \beta_3 (PX)_i^2 + \beta_4 MS_i + \beta_5 HH_i + \beta_6 ENG_i \\ & + \beta_7 IMMIG_i + \Sigma_k \beta_k COE_i + \gamma_1 MIN_i + \gamma_2 URBAN_i + \gamma_3 PT_i \\ & + \Sigma_m \gamma_m OCC_{mi} + \Sigma_n \gamma_n IND_{ni} + \varepsilon_i \end{aligned} \tag{4}$$

where:

$\ln(w)$ ≡ natural logarithm of weekly wage (in 1989 U.S. dollars)
ED ≡ years of schooling
PX ≡ potential experience (i.e., age in years – ED – 5)
MS ≡ marital status
HH ≡ householder
ENG ≡ English language ability
IMMIG ≡ born in U.S. (Canada) and residing in Canada (U.S.)
COE ≡ cohort of entry (seven periods with earliest omitted, k = 8,9,...,12)
MIN ≡ minority status
URBAN ≡ metropolitan residence status
PT ≡ part-time work status
OCC ≡ occupation of employment (six categories with one omitted, m = 4,5,...,8)
IND ≡ industry of employment (eleven categories, one omitted, n = 9,10,...,18)
ε ≡ classical stochastic error term.

Second, we estimate Equation (4) with ordinary least-squares (OLS) separately with a sample of observations from each area and for each gender. This gives us estimates of Equation (4) by gender for each of the 58 areas (116 in total; i.e., 2 × 58). Third, we partition the entire sample, irrespective of area, into two subsets: males and females. For each of these two groups, we compute the mean of each of the right-hand side variables specified in Equation (4) using the entire sample of males or females (i.e., those from all 58 areas). Using these means in the estimated Equation (4), we compute the predicted log wage for each group in each of the 58 areas. These predicted log wages constitute our estimates of the 58 area mean log wages, μ_j, for both males and females.[1] By using the entire sample of both males and females (i.e., individuals in all 58 areas) versus only individuals within each of the areas, we are able to control for interarea differences in skill mix.

An example will clarify this calculation. Say that we are interested in males in two areas: Maine and Alberta. We estimate Equation (4) for each area using the observations from that area. Net of area-specific amenity effects, there will be area-specific log wage estimates in Maine and Alberta based on both area-specific returns to skills (i.e., the β parameters) and the mean skills of male area residents. For example, *ceteris paribus,* both mean levels of education as well as differences in returns to education could result in different mean log wages between the two areas. We are concerned only with the latter effect, so we use the mean skills characteristics for all males in our sample (i.e., from all 58 areas in North America) along with the area-specific skill parameter estimates to arrive at the predicted mean log wage for each area. In this way, we are holding the skills mix constant, and allowing only the returns to skills to vary between the areas. Thus, if predicted mean log wages are higher in Alberta than in Maine, it is because of higher returns to skill in Alberta, and not due to a workforce with more education or experience, for example.

Area Returns to Skills (ϕ_j). Using Equation (3) above, we see that $\phi_j = \{\mathrm{Var}[\ln(w_{ij})^*]/\sigma^2\}^{1/2}$. To get an estimate of the variance of the log wage distribution in each area for the standardized skills distribution, $\mathrm{Var}[\ln(w_{ij})^*]$, we use Equation (4) again. This time we introduce the group-specific means, computed from the entire sample (of males or females), for each of the nonskill-related variables (i.e., those without a corresponding beta parameter) into the estimated form of Equation (4). Summing these terms with the estimated α_0 parameter yields an area-specific, constant effect on group members' log wages for each area. This constant effect does not play a role in the $\mathrm{Var}[\ln(w_{ij})^*]$.

Focusing on the skill-related terms in Equation (4), that is, those involving a beta parameter, we compute the estimated effect of these skills terms on each individual group member's log wage in area *j*. For these calculations, the entire sample of group members is used. For example, if there are 100,000 group members in the entire sample, we predict that component of the log wage of each of these 100,000 group members due to these skill-related terms by introducing each individual group member's corresponding skill-related variable values into the estimated version of Equation (4) and calculating the result. We refer to this result as the area-specific returns to skill effect for each individual. We then compute the variance of these individual area-specific returns to skill effects by group. These area-specific estimated variances are our estimates of $\mathrm{Var}[\ln(w_{ij})^*]$. There is one for each of the 58 areas for each gender. The area- specific estimate for each gender gives an estimate of the variance of the log wage distribution for the group-specific standardized skill distribution.

[1]Note that the estimated value of the α_0 parameter gives the log wage effects of area-specific factors that are common to all individual workers in the area.

In order to obtain an estimate of the variance of the standardized skill distribution for each group, we first estimate Equation (5) below with OLS for all individuals i in a group using the entire male or female sample

$$\begin{aligned}(5)\qquad \ln(w_{ij}) = {} & \alpha_0 + \Sigma_j\,\gamma_j AREA_{ij} + \beta_1 ED_{ij} + \beta_2 PX_{ij} + \beta_3 (PX)^2_{ij} + \beta_4 MS_{ij}\beta_5 HH_{ij} \\ & + \beta_6 ENG_{ij} + \beta_7 IMMIG_{ij} + \Sigma_k\,\beta_k COE_{ij} + \gamma_1 MIN_{ij} + \gamma_2 URBAN_{ij} + \gamma_3 PT_{ij} \\ & + \Sigma_m\,\gamma_m OCC_{mij} + \Sigma_n\,\gamma_n IND_{nij} + \eta_{ij}\end{aligned}$$

where *AREA* (state or province (57 areas, $j = 1,2,...,57$), with the omitted area being the reference area, and η_{ij} being a classical stochastic error term. The alpha and gamma parameters represent area-specific effects that influence the location of interarea log wage distribution but not its variance. The variance of the standardized skill distribution can be estimated for each group by, first, introducing the group means of the nonskills-related variables (based on the entire sample) into the estimated version of Equation (5) and computing the result for each group. Because group means are used, the result will not influence the variance. Second, introduce each individual group member's value for the skills-related variables (i.e., the ones premultiplied by beta parameters) into the estimated version of Equation (5) and compute the individual-specific result. These individual results provide an estimate of the returns to skill effect for each individual in each of the two groups. Finally, an estimate of σ^2 for each gender is provided by computing the variance of the individual returns for each gender.[2]

An estimate of the returns to skills parameter for each area can now be computed for each group as

$$(6)\qquad \phi_j = \{\, \mathrm{Var}[\ln(w_{ij})^*]/\sigma^2 \,\}^{1/2}$$

This parameter is, in essence, an index number that reflects the returns to skills variance in each of the areas for the standardized skill distribution—the numerator in Equation (6)—relative to the returns to skills variance for the standardized skill distribution over all 58 areas—the denominator in Equation (6). If $\phi_j > 1$ (< 1), then the area return to skills is greater than (less than) the area-wide returns to skills variance (i.e., σ^2). Because each term, $\mathrm{Var}[\ln(w_{ij})^*]$ and σ^2, are computed with the same group of individuals, the skill mix is held constant in each term, and therefore the ratio of the terms reflects differences solely in returns to skill among the 58 areas.

The differential returns to skills across areas can be computed using estimates of μ_j, ϕ_j, and σ in conjunction with Equation (1). Consider two areas $j = 1$ and 2. Ignoring stochastic errors, we see that an individual who is one standard deviation above the mean skills level of individuals in Canada and the U.S. as a whole will have the following log wages in each area

[2]Appendix Table A3 contains estimates of this equation.

$$\ln(w_{i1}) = \mu_1 + \phi_1(\sigma)$$

and

$$\ln(w_{i2}) = \mu_2 + \phi_2(\sigma)$$

The individual's log wage difference between the two areas is

$$\ln(w_{i1}) - \ln(w_{i2}) = (\mu_1 - \mu_2) + (\phi_1 - \phi_2)\sigma \tag{7}$$

Focusing on just the returns to skills component (i.e., assuming $\mu_1 = \mu_2$), we have a log wage difference between areas 1 and 2 for an individual positioned at one standard deviation above the mean of the skills distribution as equal to $(\phi_1 - \phi_2)\sigma$. The differential returns to skills measured in log wage points can be computed for all areas in this way for individuals any number of standard deviations above or below the overall mean skills level in Canada and the U.S. If the econometric estimates are obtained from some regional data other than that for the U.S. and Canada, of course, the interpretation would be the same for the computed log wage differentials across constituent areas.

3. DATA

The Canadian data are drawn from the individual file of the 1991 (3 percent) Canadian Census. The United States data are obtained by merging the 5 percent (Sample A) housing and individual records of the 1990 U.S. Census. Given the size of this file, a 5/100 subsample of the American-born was randomly generated from the sample and all Canadian-born individuals were retained. The samples were further limited to include only noninstitutionalized individuals between the ages of 25 and 64 who worked at least one week in the year prior to the census, were not self-employed, did not attend school either full- or part-time, and had at least 1000 dollars in (nominal and local currency) wage and salary income in the reference calendar year.

The income variable is the natural logarithm of the weekly wage $\ln(w)$. This is calculated by dividing the individual's yearly wage and salary income by the number of weeks worked. Because the census in either country queries respondents about wages in the year preceding the census (1989 in the United States and 1990 in Canada) and because the value of the Canadian dollar was not at par with the U.S. dollar, it was necessary to adjust the weekly wage variable. To do so, Canadian weekly wages in 1990 were first deflated by the year-over-year inflation rate (4.8 percent) and then deflated to U.S. dollars using the 1989 exchange rate of 1.184.

The years of education variable *ED* was coded to equal the number of years corresponding to the highest level of education completed. For example, high school graduation or GED was coded to equal 12, some post-secondary education or an associate degree was coded as 14 years of education, while a bachelor's degree was coded as 16, a master's or professional degree as 18 years, and a doctorate as 20. Potential experience *PX* was calculated using the familiar

Mincerian proxy (i.e., age in years – *ED* – 5). In both the U.S. and Canadian data the age variable is continuous, but because the variable ED was derived from discrete intervals, *PX* was marginally negative in a few cases and was therefore bottom-coded to zero. The marital status variable *MS* was coded to unity if the respondent said that he or she was married, and zero otherwise. If this respondent was a household head or householder, the variable *HH* was coded to unity.

Immigrants from either country are denoted by a dummy variable *IMMIG* set equal to unity. Immigrants are also classified by their period of immigration by a number of cohort of entry *COE* variables. These variables are slightly different for each country, although there are seven separate dummy variables in each case. In the Canadian data, these entry cohorts are before 1961, followed by six five-year entry cohorts beginning with 1961–1965 and ending with 1986–1990. For the United States, these seven dummies are roughly for the same time periods—before 1960 followed by six separate five-year cohorts beginning with 1960–1964 and ending with 1985–1989. These seven entry cohorts in either country are denoted as *YPRE61*, *Y61*, . . . , *Y86*.

MIN is a dummy variable coded to unity if the individual is a visible minority. *URBAN* is a dummy variable reflecting residency in an urban area. These definitions do differ somewhat between the two countries. The variable *PT* is a dummy that was set to equal one if the respondent worked part time (i.e., less than 30 hours per week). This variable existed in the Canadian data, and was derived from the U.S. data to be consistent with the Canadian definition. Finally, the variables *OCC* and *IND* are occupation and industry variables, respectively. These were coded to be as consistent as possible across the two censuses.

The final Canadian male sample consists of 114,934 individuals, compared to a total of 105,489 male individuals in the American data. For females, the Canadian data include 98,663 individuals while the U.S. data contain 94,425. Appendix Tables A1 and A2 present summary statistics for the male and female samples.

4. ESTIMATION RESULTS

Males

Table 1 presents the estimates of μ_j (the standardized mean weekly wage) and ϕ_j (the standardized index of returns to skills), and area-specific log wage differentials for individuals at various point in the skills distribution (i.e., one, two, and three standard deviations above the mean) for males in each of the 58 areas, as well as weighted averages for the 10 Canadian provinces, and the lower 48 states. If $\phi_j > 1$ (< 1), this indicates that that area has a returns to skills distribution more (less) dispersed than the weighted average for all 58 areas. Three noteworthy points are immediately apparent from these estimates. First, the average estimated log weekly wage in the United States is about the same as that in Canada (once the latter is adjusted for price levels and exchange rate

TABLE 1: μ, φ, and Returns to Skill, States (1989) and Provinces (1990), Males

		μ_j		ϕ_j	+ 1σ	+ 2σ	+ 3σ	
	Number of Observations	Value	Rank	Values				Rank
Delaware	322	6.341	7	1.477	0.318	0.637	0.955	1
North Dakota	241	6.054	46	1.468	0.317	0.633	0.950	2
Maryland	2420	6.353	5	1.344	0.290	0.580	0.869	3
Utah	682	6.239	15	1.310	0.282	0.565	0.847	4
Wyoming	197	6.147	32	1.297	0.280	0.559	0.839	5
Texas	6900	6.194	24	1.285	0.277	0.554	0.831	6
Georgia	2965	6.202	22	1.284	0.277	0.554	0.830	7
New York	7157	6.344	6	1.268	0.273	0.547	0.820	8
New Mexico	596	6.159	29	1.265	0.273	0.545	0.818	9
New Jersey	3407	6.508	1	1.251	0.270	0.539	0.809	10
Tennessee	2241	6.069	45	1.249	0.269	0.539	0.808	11
Ohio	4839	6.228	16	1.248	0.269	0.538	0.807	12
California	10747	6.394	4	1.240	0.267	0.535	0.802	13
Michigan	4226	6.297	9	1.235	0.266	0.533	0.799	14
Kentucky	1526	6.145	33	1.230	0.265	0.531	0.796	15
Connecticut	1595	6.508	2	1.222	0.264	0.527	0.791	16
Iowa	1136	6.084	43	1.222	0.264	0.527	0.791	17
Mississippi	1004	6.082	44	1.222	0.263	0.527	0.790	18
Massachusetts	3089	6.404	3	1.214	0.262	0.524	0.785	19
Louisiana	1619	6.163	27	1.214	0.262	0.523	0.785	20
North Carolina	3042	6.118	37	1.207	0.260	0.520	0.781	21
Virginia	2844	6.264	14	1.206	0.260	0.520	0.780	22
Missouri	2260	6.160	28	1.198	0.258	0.517	0.775	23
Indiana	2502	6.196	23	1.186	0.256	0.511	0.767	24
Colorado	1600	6.221	18	1.175	0.253	0.507	0.760	25
Illinois	4920	6.310	8	1.172	0.253	0.505	0.758	26
South Dakota	231	6.004	47	1.168	0.252	0.504	0.756	27
South Carolina	1494	6.131	35	1.158	0.250	0.499	0.749	28
Rhode Island	394	6.272	13	1.153	0.249	0.497	0.746	29
Alabama	1628	6.097	40	1.120	0.242	0.483	0.725	30
West Virginia	731	6.102	38	1.106	0.238	0.477	0.715	31
Arizona	1516	6.218	20	1.100	0.237	0.474	0.712	32
Washington	2510	6.281	11	1.089	0.235	0.470	0.704	33
Florida	4935	6.177	25	1.085	0.234	0.468	0.702	34
Pennsylvania	5446	6.203	21	1.075	0.232	0.463	0.695	35
Montana	339	6.124	36	1.073	0.231	0.463	0.694	36
New Hampshire	650	6.285	10	1.051	0.227	0.453	0.680	37
Oklahoma	1212	6.112	38	1.044	0.225	0.450	0.676	38
Arkansas	892	6.003	48	1.038	0.224	0.448	0.671	39
Kansas	1072	6.149	30	1.006	0.217	0.434	0.651	40
Wisconsin	2183	6.177	26	0.979	0.211	0.422	0.633	41
Vermont	320	6.133	32	0.966	0.208	0.417	0.625	42
Oregon	1337	6.222	17	0.953	0.206	0.411	0.617	43
Maine	733	6.148	31	0.946	0.204	0.408	0.612	44
Nevada	641	6.273	12	0.936	0.202	0.404	0.605	45
Nebraska	647	6.096	41	0.931	0.201	0.402	0.602	46
Minnesota	2111	6.221	19	0.915	0.197	0.394	0.592	47

	Number of Observations	μ_j Value	Rank	ϕ_j	+ 1σ	+ 2σ	+ 3σ	Rank
				Values				
Idaho	404	6.089	42	0.883	0.190	0.381	0.571	48
U.S. Average	*105489*	*6.250*		*1.184*	*0.255*	*0.510*	*0.766*	
Ontario	38579	6.333	2	0.881	0.190	0.380	0.570	1
Manitoba	4495	6.161	8	0.878	0.189	0.379	0.568	2
Alberta	10960	6.272	3	0.853	0.184	0.368	0.552	3
Saskatchewan	3851	6.158	9	0.851	0.183	0.367	0.550	4
Quebec	32435	6.237	4	0.819	0.177	0.353	0.530	5
Newfoundland	2909	6.170	7	0.780	0.168	0.336	0.504	6
New Brunswick	3731	6.189	6	0.749	0.161	0.323	0.484	7
British Columbia	12879	6.350	1	0.738	0.159	0.318	0.477	8
Nova Scotia	4501	6.191	5	0.701	0.151	0.302	0.454	9
Prince Edward Island	594	6.056	10	0.615	0.133	0.265	0.398	10
Canada Average	*114934*	*6.274*		*0.829*	*0.179*	*0.357*	*0.536*	
Canada and U.S. Average	*220423*	*6.262*		*0.999*	*0.215*	*0.431*	*0.646*	

Note: σ = 0.2156 and is the value of the standard deviation of the standardized skill distribution.

differentials). In the U.S. the average estimated log wage is 6.250, compared to the Canadian average of 6.274.

There are also interregional differences in these estimates within each country. In the United States, the five highest mean log weekly wages are in the Northeast (New Jersey, Connecticut, Massachusetts, and Maryland) along with California. By contrast, three of the five states with the lowest mean log weekly wages are in the South (Arkansas, Tennessee, and Mississippi) along with the Dakotas. In Canada, the three "have" provinces (Alberta, British Columbia, and Ontario) have the highest mean log weekly wages, while Manitoba, Saskatchewan, and Prince Edward Island have the lowest.

Second, the value of ϕ is higher in the United States than in Canada: an average of 1.184 in the United States compared to 0.829 in Canada. This result is consistent with the sizeable literature on comparative earnings and income distributions.[3] In fact, in all cases, the ϕs are higher for each of the U.S. states than for any of the Canadian provinces.[4]

Third, an individual's place in the standardized skills distribution, along with the relative returns to skills in each area, interact to determine an individual's position above or below the mean area-specific log weekly wage. The final four columns of Table 1 illustrate the returns to skills for individuals one, two, and three standard deviations above the mean of the standardized skills

[3]Recent work includes Blackburn and Bloom (1993), Gottschalk and Smeeding (1997), and Richardson (1997).

[4]The interregional sample correlations between μ and ϕ are 0.157 (males) and –0.158 (females). Neither is statistically significant at even the 20 percent level for a two-tailed test. In contrast, the interregional sample correlation between male and female μ values is 0.853 (p-value = 0.00), and that between male and female ϕ values is 0.648 (p-value = 0.00).

distribution, as well as the within-country rankings of each of these areas. As an illustration, an individual with skills one standard deviation above the mean, and residing in the Delaware, would have log weekly wages that are 0.318 log points higher than an individual with mean skills. If another individual is three standard deviations above the mean, this differential increases to 0.955 log points. Observationally equivalent males in Idaho, the state with the most compressed distribution, would earn only 0.190 and 0.571 log points, respectively, above mean area log wages.

Following Equation (7), these results imply that *ceteris paribus*, a person with skills at the mean of the standardized distribution, and who moves to Delaware from Idaho, would see his weekly wages increase by 0.252 log points (i.e., 6.341 – 6.089). Another individual making the same relocation decision, but at one standard deviation above the mean would increase his log wages by a further 0.128 log points (i.e., 0.318 – 0.190). An individual at the third standard deviation above the mean of the same distribution would earn weekly wages 0.384 log points higher than the individual at the mean (i.e., 0.955 – 0.571).

Females

Table 2 presents comparable estimates for females. The results are largely the same as for males. Mean log wages in Canada and the United States are different (means of 5.799 versus 5.671, respectively). Again, the average value of ϕ in the United States is much higher than that in Canada (1.254 compared to 0.755). Only three states have ϕ values below the highest provincial value. The female data for mean log weekly wages share a similar pattern to the male data in terms of regional differences within each country. In fact, four of the five states with the highest mean log wages are identical: three in the Northeast (Connecticut, New Jersey, and Massachusetts) along with California. The lowest estimates again include the Dakotas, but also Montana, Oklahoma, and Iowa. Similarly, in Canada, the provinces with the highest values are Ontario, Alberta and British Columbia, while Prince Edward Island, Nova Scotia, and New Brunswick are at the bottom.

5. CONCLUSIONS

In part following previous studies in the spatial general equilibrium literature on interregional wage variations, we present a method for identifying both the mean of a region's wage distribution and the region's returns to skills parameter. The latter is related to the variance of the region's wage distribution. A key aspect of our method is the development of a standardized skill distribution that permits us to isolate regional variations in both the skills mix and returns to skills components of interregional wage variations. Using individual micro data from the U.S. and Canadian censuses, we estimate returns to skills parameters and the means of wage distributions for 48 U.S. states and 10 Canadian provinces for the period 1989–1990. In general, we find that although mean log weekly wages in the United States and Canada are similar,

TABLE 2: μ, φ, and Returns to Skill, States (1989) and Provinces (1990), Females

		μ_j		ϕ_j	+ 1σ	+ 2σ	+ 3σ	
	Number of Observations	Value	Rank		Values			Rank
West Virginia	527	5.508	40	1.879	0.269	0.537	0.806	1
Vermont	268	5.510	39	1.686	0.241	0.482	0.723	2
New Mexico	514	5.589	25	1.584	0.227	0.453	0.680	3
Georgia	2809	5.639	18	1.570	0.224	0.449	0.673	4
New York	6607	5.776	6	1.542	0.220	0.441	0.661	5
Texas	5989	5.626	19	1.488	0.213	0.426	0.638	6
Connecticut	1546	5.920	1	1.480	0.212	0.423	0.635	7
Maryland	2174	5.765	7	1.469	0.210	0.420	0.630	8
Kentucky	1298	5.506	41	1.459	0.209	0.417	0.626	9
Mississippi	936	5.531	36	1.419	0.203	0.406	0.609	10
Delaware	264	5.799	5	1.405	0.201	0.402	0.603	11
Virginia	2569	5.697	12	1.400	0.200	0.400	0.601	12
Louisiana	1338	5.523	37	1.391	0.199	0.398	0.597	13
Idaho	383	5.539	32	1.350	0.193	0.386	0.579	14
Alabama	1500	5.537	34	1.286	0.184	0.368	0.552	15
New Jersey	3146	5.900	2	1.283	0.183	0.367	0.550	16
Florida	4947	5.675	14	1.279	0.183	0.366	0.549	17
Maine	658	5.588	26	1.277	0.183	0.365	0.548	18
Pennsylvania	4472	5.606	23	1.271	0.182	0.364	0.545	19
Michigan	3670	5.688	13	1.255	0.179	0.359	0.538	20
North Carolina	2974	5.613	22	1.247	0.178	0.357	0.535	21
Arizona	1289	5.648	15	1.227	0.175	0.351	0.526	22
Wyoming	175	5.500	42	1.198	0.171	0.343	0.514	23
California	9365	5.867	3	1.186	0.170	0.339	0.509	24
New Hampshire	626	5.726	9	1.181	0.169	0.338	0.507	25
Missouri	2002	5.544	31	1.176	0.168	0.336	0.504	26
South Dakota	267	5.438	46	1.163	0.166	0.333	0.499	27
North Dakota	245	5.412	48	1.147	0.164	0.328	0.492	28
Tennessee	1973	5.562	29	1.140	0.163	0.326	0.489	29
Illinois	4319	5.644	16	1.139	0.163	0.326	0.488	30
Utah	611	5.539	33	1.130	0.162	0.323	0.485	31
Colorado	1395	5.623	20	1.130	0.162	0.323	0.485	32
Wisconsin	2043	5.580	27	1.120	0.160	0.320	0.481	33
Ohio	4209	5.600	24	1.117	0.160	0.319	0.479	34
Rhode Island	414	5.749	8	1.110	0.159	0.317	0.476	35
Minnesota	1796	5.639	17	1.100	0.157	0.314	0.472	36
Oregon	1104	5.620	21	1.087	0.155	0.311	0.466	37
Oklahoma	1162	5.485	45	1.083	0.155	0.310	0.465	38
South Carolina	1378	5.536	35	1.073	0.153	0.307	0.460	39
Washington	2125	5.724	10	1.063	0.152	0.304	0.456	40
Arkansas	835	5.493	43	1.051	0.150	0.301	0.451	41
Massachusetts	2791	5.819	4	1.031	0.147	0.295	0.442	42
Indiana	2331	5.553	30	0.981	0.140	0.280	0.421	43
Kansas	922	5.579	28	0.976	0.140	0.279	0.419	44
Nevada	526	5.700	11	0.907	0.130	0.260	0.389	45
Nebraska	587	5.516	38	0.818	0.117	0.234	0.351	46
Montana	288	5.426	47	0.815	0.117	0.233	0.350	47

	Number of Observations	μ_j Value	μ_j Rank	ϕ_j	+ 1σ	+ 2σ	+ 3σ	Rank
				Values				
Iowa	1070	5.487	44	0.682	0.098	0.195	0.293	48
U.S. Average	*94425*	*5.671*		*1.254*	*0.179*	*0.359*	*0.538*	
Quebec	26523	5.778	4	0.823	0.118	0.235	0.353	1
Ontario	34130	5.853	1	0.818	0.117	0.234	0.351	2
Nova Scotia	3739	5.688	9	0.779	0.111	0.223	0.334	3
Alberta	9634	5.790	3	0.698	0.100	0.200	0.299	4
Manitoba	4056	5.718	6	0.653	0.093	0.187	0.280	5
British Columbia	10878	5.851	2	0.650	0.093	0.186	0.279	6
New Brunswick	3031	5.654	10	0.620	0.089	0.177	0.266	7
Saskatchewan	3754	5.707	7	0.566	0.081	0.162	0.243	8
Prince Edward Island	573	5.698	8	0.545	0.078	0.156	0.234	9
Newfoundland	2345	5.727	5	0.480	0.069	0.137	0.206	10
Canada Average	*98663*	*5.799*		*0.755*	*0.108*	*0.216*	*0.324*	
Canada and U.S. Average	*193376*	*5.737*		*0.999*	*0.143*	*0.286*	*0.429*	

Note: σ = 0.1430 and is the value of the standard deviation of the standardized skill distribution.

almost all states in the United States have greater returns to skills than Canadian provinces. This implies that the regional returns to skills are systematically higher in the U.S. than in Canada. These results hold for both the male and female subsamples.

Although a thorough examination of the reasons for and implications of these interregional differences is beyond the scope of this methodological paper, theoretical reasons for these differences do exist. These could include regional differences in complementary inputs (i.e., physical capital, human capital, and natural resources), urbanization and other agglomeration effects, labor supply effects of social programs on labor force participation (which could have an effect on inclusion in our sample), or even cultural differences may explain part or even all of these differences. These would all appear to be fruitful areas for future research, as would the implications of the estimated interregional variations in μ and ϕ for interregional migration, human capital-based regional growth, and interregional convergence and compensating differentials.

REFERENCES

Beeson, Patricia E. 1991. "Amenities and Regional Differences in Returns to Worker Characteristics," *Journal of Urban Economics,* 30, 224–231.

Beeson, Patricia E. and Randall W. Eberts. 1989. "Identifying Productivity and Amenity Effects in Interurban Wage Differentials," *The Review of Economics and Statistics,* 71, 443–452.

Blackburn, McKinley L. and David E. Bloom. 1993. "The Distribution of Family Income: Measuring and Explaining Changes in the 1980s for Canada and the United States," in D. Card and R. B. Freeman (eds.), *Small Differences that Matter: Labor Markets and Income Maintenance in Canada and the United States*. Chicago: University of Chicago Press, pp. 233–265.

Borjas, George J., Stephen G. Bronars, and Stephen J. Trejo. 1992. "Self-Selection and Internal Migration in the United States," *Journal of Urban Economics,* 32, 159–187.

Blomquist, Glenn C., Mark C. Berger, and John P. Hoehn. 1988. "New Estimates of Quality of Life in Urban Areas," *American Economic Review,* 78, 89–107.

Dickie, Mark and Shelby Gerking. 1987. "Interregional Wage Differentials: An Equilibrium Perspective," *Journal of Regional Science,* 27, 571–585.

———. 1998. "Interregional Wage Disparities, Relocation Costs, and Labor Mobility in Canada," *Journal of Regional Science,* 38, 61–87.

Gottschalk, Peter and Timothy M. Smeeding. 1997. "Cross-National Comparisons of Earnings and Income Inequality," *Journal of Economic Literature,* 35, 633–687.

Greenwood, Michael J., Gary L. Hunt, Dan S. Rickman, and George I. Treyz. 1991. "Migration, Regional Equilibrium, and the Estimation of Compensating Differentials," *American Economic Review,* 81, 1382–1390.

Hanushek, Eric A. 1973. "Regional Differences in the Structure of Earnings," *Review of Economics and Statistics,* 55, 204–213.

Hunt, Gary L. and Richard E. Mueller. 2000. "Canadian Immigration to the U.S. 1985–1990: Estimates from a Roy Selection Model of Differences in Returns to Skill," unpublished manuscript.

Rauch, James E. 1993. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities," *Journal of Urban Economics,* 28, 71–86.

Richardson, David H. 1997. "Changes in the Distribution of Wages in Canada, 1981–1992," *Canadian Journal of Economics,* 30, 622–643.

Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life," *Journal of Political Economy,* 90, 1257–1278.

———. 1988. "Wages, Rents, and Amenities: Differences Among Workers and Regions," *Economic Inquiry,* 26, 23–41.

APPENDIX

TABLE A1: Selected Summary Statistics, Males (n = 220,423)

Variable Name	Mean	Standard Deviation	Minimum	Maximum
Natural Log of the Weekly Wage ($ln(w)$)	6.274	0.641	2.741	11.490
Years of Schooling (*ED*)	12.995	2.613	5	20
Potential Experience (*PX*)	22.401	10.996	0	54
Married (*MS*)	0.752	0.432	0	1
Householder (*HH*)	0.812	0.391	0	1
English Language Ability (*ENG*)	0.923	0.267	0	1
Immigrant (*IMMIG*)	0.036	0.187	0	1
Cohort of Entry (*COE*)				
YPRE61	0.012	0.111	0	1
Y61	0.006	0.079	0	1
Y66	0.005	0.071	0	1
Y71	0.003	0.058	0	1
Y76	0.003	0.052	0	1
Y81	0.003	0.052	0	1
Y86	0.003	0.059	0	1
Minority (*MIN*)	0.054	0.227	0	1
Metropolitan Resident (*URBAN*)	0.674	0.469	0	1
Part-time Work (*PT*)	0.033	0.179	0	1
Occupation Code 1 (*OCC1*)	0.137	0.344	0	1
Occupation Code 2 (*OCC2*)	0.119	0.323	0	1
Occupation Code 3 (*OCC3*)	0.049	0.216	0	1
Occupation Code 4 (*OCC4*)	0.171	0.377	0	1
Occupation Code 5 (*OCC5*)	0.099	0.299	0	1
Occupation Code 6 (*OCC6*)	0.425	0.494	0	1
Industry of Employment Code 1 (*IND1*)	0.051	0.219	0	1
Industry of Employment Code 2 (*IND2*)	0.100	0.300	0	1
Industry of Employment Code 3 (*IND3*)	0.238	0.426	0	1
Industry of Employment Code 4 (*IND4*)	0.122	0.328	0	1
Industry of Employment Code 5 (*IND5*)	0.061	0.240	0	1
Industry of Employment Code 6 (*IND6*)	0.097	0.296	0	1
Industry of Employment Code 7 (*IND7*)	0.045	0.207	0	1
Industry of Employment Code 8 (*IND8*)	0.044	0.204	0	1
Industry of Employment Code 9 (*IND9*)	0.049	0.215	0	1
Industry of Employment Code 10 (*IND10*)	0.107	0.309	0	1
Industry of Employment Code 11 (*IND11*)	0.086	0.281	0	1
Standardized Skills	0.000	0.216	–0.781	0.753

TABLE A2: Selected Summary Statistics, Females (n = 193,088)

Variable Name	Mean	Standard Deviation	Minimum	Maximum
Natural log of the weekly wage (*ln*(*w*))	5.742	0.710	2.741	10.786
Years of Schooling (*ED*)	13.063	2.297	5	20
Potential Experience (*PX*)	22.042	10.830	0	54
Married (*MS*)	0.704	0.457	0	1
Householder (*HH*)	0.293	0.455	0	1
English Language Ability (*ENG*)	0.919	0.272	0	1
Immigrant (*IMMIG*)	0.044	0.205	0	1
Cohort of Entry (*COE*)				
YPRE61	0.016	0.125	0	1
Y61	0.008	0.088	0	1
Y66	0.006	0.078	0	1
Y71	0.004	0.065	0	1
Y76	0.004	0.059	0	1
Y81	0.003	0.056	0	1
Y86	0.003	0.057	0	1
Minority (*MIN*)	0.066	0.248	0	1
Metropolitan Resident (*URBAN*)	0.691	0.462	0	1
Part-time Work (*PT*)	0.196	0.397	0	1
Occupation Code 1 (*OCC1*)	0.093	0.291	0	1
Occupation Code 2 (*OCC2*)	0.175	0.380	0	1
Occupation Code 3 (*OCC3*)	0.049	0.215	0	1
Occupation Code 4 (*OCC4*)	0.257	0.437	0	1
Occupation Code 5 (*OCC5*)	0.325	0.468	0	1
Occupation Code 6 (*OCC6*)	0.102	0.302	0	1
Industry of Employment Code 1 (*IND1*)	0.019	0.136	0	1
Industry of Employment Code 2 (*IND2*)	0.016	0.126	0	1
Industry of Employment Code 3 (*IND3*)	0.123	0.328	0	1
Industry of Employment Code 4 (*IND4*)	0.053	0.223	0	1
Industry of Employment Code 5 (*IND5*)	0.030	0.171	0	1
Industry of Employment Code 6 (*IND6*)	0.142	0.349	0	1
Industry of Employment Code 7 (*IND7*)	0.090	0.286	0	1
Industry of Employment Code 8 (*IND8*)	0.043	0.202	0	1
Industry of Employment Code 9 (*IND9*)	0.085	0.280	0	1
Industry of Employment Code 10 (*IND10*)	0.330	0.470	0	1
Industry of Employment Code 11 (*IND11*)	0.070	0.255	0	1
Standardized Skills	0.000	0.143	–0.581	0.565

TABLE A3: Estimates of Equation (5), Males and Females

	Males		Females	
Variable	Coefficient	Standard Error	Coefficient	Standard Error
Years of Schooling (*ED*)	0.0636	0.0006	0.0617	0.0008
Potential Experience (*PX*)	0.0294	0.0005	0.0129	0.0005
Potential Experience Squared (PX^2)	–0.0004	0.0000	–0.0002	0.0000
Married (*MS*)	0.1445	0.0030	0.0148	0.0037
Householder (*HH*)	0.1385	0.0033	0.0833	0.0037
English Language Ability (*ENG*)	0.0003	0.0059	0.0181	0.0068
Immigrant (*IMMIG*)	0.0481	0.0107	0.0185	0.0107
Cohort of Entry (*COE*)				
Y61	0.0251	0.0180	0.0108	0.0181
Y66	0.0047	0.0195	0.0154	0.0196
Y71	–0.0102	0.0228	–0.0113	0.0226
Y76	0.0830	0.0248	–0.0154	0.0243
Y81	0.0304	0.0247	0.0033	0.0254
Y86	0.0908	0.0224	0.0142	0.0251
Minority (*MIN*)	–0.1578	0.0054	–0.0174	0.0056
Metropolitan Resident (*URBAN*)	–0.0415	0.0065	0.1252	0.0141
Part-time Work (*PT*)	–0.0215	0.0059	0.1829	0.0103
Occupation Code 2 (*OCC2*)	0.0019	0.0063	0.2633	0.0112
Occupation Code 3 (*OCC3*)	–0.0866	0.0072	0.1381	0.0123
Occupation Code 4 (*OCC4*)	–0.2562	0.0067	–0.0692	0.0105
Occupation Code 5 (*OCC5*)	–0.0534	0.0080	0.1520	0.0107
Occupation Code 6 (*OCC6*)	–0.1663	0.0079	0.0642	0.0116
Industry of Employment Code 2 (*IND2*)	–0.3507	0.0077	–0.1651	0.0109
Industry of Employment Code 3 (*IND3*)	–0.2322	0.0068	0.0693	0.0101
Industry of Employment Code 4 (*IND4*)	–0.0927	0.0068	0.1756	0.0109
Industry of Employment Code 5 (*IND5*)	–0.0838	0.0049	0.0353	0.0057
Industry of Employment Code 6 (*IND6*)	–0.1763	0.0062	–0.1080	0.0075
Industry of Employment Code 7 (*IND7*)	–0.2675	0.0044	–0.3453	0.0053
Industry of Employment Code 8 (*IND8*)	–0.2893	0.0050	–0.2619	0.0050
Industry of Employment Code 9 (*IND9*)	–0.2793	0.0041	–0.3182	0.0067
Industry of Employment Code 10 (*IND10*)	0.0987	0.0028	0.1163	0.0032
Industry of Employment Code 11 (*IND11*)	–0.5580	0.0066	–0.6149	0.0034
Constant	5.0008	0.0259	4.9836	0.0296
R^2	0.2700	0.3500		
Number of Observations	220,423	193,088		

Note: Estimates of area-specific fixed-effects have not been reported.