



DEPARTAMENTO DE ECONOMÍA

**SDT 321**

## **THE PLACE PREMIUM: WAGE DIFFERENCES FOR IDENTICAL WORKERS ACROSS THE US BORDER**

Michael Clemens  
Claudio Montenegro  
Lant Pritchett

Santiago, julio 2010

La serie de Documentos de Trabajo (SDT) del Departamento de Economía de la Universidad de Chile en versión PDF puede descargarse en la dirección electrónica [www.econ.uchile.cl/SDT](http://www.econ.uchile.cl/SDT). Para contactar al editor ejecutivo de SDT remitirse a [sdt@econ.uchile.cl](mailto:sdt@econ.uchile.cl)

## **The Place Premium: Wage Differences for Identical Workers across the US Border**

Michael A. Clemens

Center for Global  
Development

Claudio E. Montenegro

World Bank and  
Dept. of Economics,  
Universidad de Chile

Lant Pritchett

Harvard Kennedy School  
and Center for Global  
Development

*Abstract:* We estimate the “place premium”—the wage gain that accrues to foreign workers who arrive to work in the United States. First, we estimate the predicted, purchasing-power adjusted wages of people inside and outside the United States who are otherwise observably identical—with the same country of birth, country of education, years of education, work experience, sex, and rural or urban residence. We use new and uniquely rich micro-data on the wages and characteristics of over two million individual formal-sector wage-earners in 43 countries (including the US). Second, we examine the extent to which these wage ratios for observably equivalent workers may overstate the gains to a marginal mover because movers may be positively selected on unobservable productivity in their home country. New evidence for nine of the countries, combined with a range of existing evidence, suggests that this overstatement can be significant, but is typically modest in magnitude. Third, we estimate the degree to which policy barriers to labor movement in and of themselves sustain the place premium, by bounding the premia observed under self-selected migration alone. Finally, we show that the policy-induced portion of the place premium in wages represents one of the largest remaining price distortions in any global market; is much larger than wage discrimination in spatially integrated markets; and makes labor mobility capable of reducing households’ poverty at the margin by much more than any known *in situ* intervention.

**Keywords:** migration, wage discrimination, price distortions, policy barriers, place premium, poverty.

**JEL Codes** F22, J61, J71, O15.

We are grateful to Indermit S. Gill and his team at the World Development Report of the World Bank who built and allowed use of the database. We thank Samuel Bazzi and Paolo Abarcar for excellent research assistance. We received helpful comments from Christopher Blattman, William Easterly, David Lindauer, David McKenzie, and seminar participants at Yale University, Harvard Kennedy School, the Brookings Institution, the LACEA annual meetings, and the Center for Global Development. This work was partially supported by generous grants from the John D. and Catherine T. MacArthur Foundation and from AusAID. The paper represents the views of the authors alone and not necessarily those of their employers or funders.

## 1 Introduction

Two facts are obvious to even the most casual traveler in the contemporary, supposedly globalized, world. Every (legal) traveler's very first experience in every country is an encounter with the agent of the state responsible for enforcing that country's restrictions on the international movement of people—especially workers. Thus the single most obvious fact to a global traveler is the enforcement by every country of a complex, quantity-based set of impediments to the movement of labor. The second most obvious fact, at least to any traveler outside the rich world, is the enormous differences across countries in wages and standards of living.

We examine the connection between those two facts: barriers to labor mobility, and international wage gaps. Our goal is an estimate of magnitude of the “place premium” to working the US—the difference in the real (consumption deflated) wages of workers of equal intrinsic productivity on opposite sides of the US border—that is sustained by restrictions on labor mobility. This ambitious goal requires three steps.

First, we use a unique harmonized database on the purchasing power-adjusted wages and other traits of over two million workers in 42 countries plus the United States. This allows us to predict the wages of observably identical workers on either side of the US border for people from each of those countries. Crucially, the US census data identify each individual's country of birth and, for the foreign-born, their year of arrival in the US. Hence our definition of “observably identical” allows us to compare not only people with the same standard traits (years of schooling, age, sex, rural/urban residence) but different nationality, but to directly compare workers born and educated in the same country. This implicitly controls for nationality-specific characteristics that affect productivity in the US (e.g. culture, language) and the quality and relevance of a country's schooling to US labor market outcomes. For instance, in our preferred econometric specification, a Peruvian-born, Peruvian-educated, 35 year-old urban male formal sector wage-worker with 9 years of schooling earns an average of \$1,714 per month working in the United States but the average person with these observable traits

earns P\$452 (P\$ are purchasing power parity adjusted US dollars) working in Peru. The wage ratio,  $R_o$  (where the subscript signifies “observably identical”), is 3.8.

For each of the 42 countries, we estimate  $R_o$ : the ratio of wages earned by workers in the United States to wages earned by observably identical workers abroad.<sup>1</sup> The typical wage ratio is very large: Peru’s ratio  $R_o$  of 3.8 is near the median of 4.11. And this ratio varies greatly across countries: Estimates of  $R_o$  range from 2.0 for the Dominican Republic to 15.5 for Yemen. The 75<sup>th</sup> percentile of  $R_o$  is 6.5, while the 25<sup>th</sup> percentile is 2.8.

The second part of the paper grapples with the issue, common to all attempts to measure wage discrimination, of selectivity. What we need is an estimate of the counter-factual: the wages of the workers in the US had they remained in their home country. No matter how many individual traits are controlled for, wage differentials for observably equivalent workers are not exact estimates of wage differentials between workers of equal intrinsic productivity. While our estimates of  $R_o$  account for selection on observables, some part of the wage gap between foreign-born workers in the US and observably identical workers abroad may be due to selection on unobservable determinants of productivity. This could be due to self-selection (e.g. those with more unobservable “pluck” move) and selection by migration policy or US employers (which might select those among a group of observably identical workers with the highest productivity). We need a method to estimate  $R_e$  (where the subscript denotes “equivalent”), the cross-border wage ratio for an observably *and* unobservably identical worker—the same person.  $R_e$ , not  $R_o$ , is the place premium.

We therefore present new evidence about the degree of migrant selection on *unobservable* determinants of wages, for nine of the 42 countries where existing data allow the calculation. This permits us to roughly estimate  $R_e$  for nine countries. For example, the evidence suggests that emigrants from Peru come from somewhere around

---

<sup>1</sup> Rosenzweig (2006) also estimates gains to movement from differences in the “skill price”—the price of equivalently skilled labor in different labor markets—using observed wage changes of the same individuals from the US New Immigrant Survey. His analysis focuses more on higher-skill workers within a joint model of mobility decisions and skill acquisition (e.g. the decision to seek higher education in the US) whereas in this paper we focus on lower-skill workers using individual-level survey data across countries.

the 69<sup>th</sup> percentile of the distribution of unobserved wage determinants. This implies that an observably *and* unobservably identical prime-age urban formal-sector male Peruvian with nine years of Peruvian schooling earns about 2.6 times as much in the US as in Peru. This is one of several cases where our estimates of  $R_o$  for moderately skilled workers exceed the place premium  $R_e$ . Even after this correction, however, the estimates of  $R_e$  are very large—including 3.5 for the Philippines and 7.8 for Haiti. These estimates of  $R_e$  are “marginal” in two distinct senses: (i) it is the effect on the wage of the next person who would arrive after a small relaxation of the migration barrier—*not* the effect of moving the average person chosen at random from the sending country; and (ii) it is the marginal effect given a small relaxation of current restrictions—*not* the general equilibrium wages under fully open borders which involves considerations of how mobile labor would affect capital accumulation and Total Factor Productivity (Klein and Ventura 2004).

We corroborate these findings with a range of other evidence, both microeconomic and macroeconomic, that bears on the degree of migrant selection. These calculations yield the remarkably consistent result that selection of moderate skill movers on unobservable home-country wage determinants results in a ratio of  $R_o/R_e$  (the ratio of the observably identical to “equal productivity” wage ratios) of around 1.2, varying from 1.0 (no bias at all) to about 1.5 (substantial positive selection) for different countries. Even adjusting our estimates of  $R_o$  for the 42 countries for reasonable estimates of the degree of selection, the resulting estimates of  $R_e$  remain very large for most countries. For instance, even if one assumed that the highest degree of positive selection on unobservable wage determinants seen for *any* country applies to *all* 42 countries,<sup>2</sup>  $R_e$  would still exceed 3 in 20 out of 42 countries and would still exceed 2 (a doubling of wages from crossing the border) in all but four.

Once we have estimates, direct or indirect, of  $R_e$ , the third step is to ask how much of the observed differences in wages of equal intrinsic productivity workers across the border is due to policy barriers to labor mobility and how much could be attributed to “natural”

---

<sup>2</sup> That is, assuming that the mean migrant’s counterfactual home-country wage would equal the 70<sup>th</sup> percentile of wages for an observably identical worker (35 year-old urban male with 9 years of education).

barriers that would cause equilibrium  $R_e > 1$  even in “borderless” labor markets. Workers might require a compensating differential to bear the costs—broadly considered—of moving to a new land. These include the difficulty of learning a new language, being away from one’s family, and entering new social networks, as well as the direct cost of travel. Workers might also be credit-constrained and have difficulty financing the move. We estimate the wage differentials consistent with free mobility using data from a variety of contemporary and historical situations with legally integrated, but spatially separated and cultural distinct labor markets. These data suggest that real wage ratios higher than 1.5 to 1.8 are unlikely to be sustained by natural barriers alone. Wage ratios higher than this are consistent with either substantial labor mobility or policy induced barriers.

The paper concludes by relating our results to three separate literatures on border-induced price wedges, wage discrimination, and the marginal impacts of antipoverty policies. To make a crude and conservative estimate, suppose we begin with the median estimated wage gap for an observably identical worker of about P\$15,000 per year. Adjusting this figure both for reasonable estimates of migrant selection on unobservables and for compensating differentials, this suggests that existing border distortions produce an available welfare gain to a marginal moderate-skill mover from a typical developing country of around P\$10,000 year. This is a massive cross-border price wedge and dwarfs the welfare gains from liberalizations in other markets. Legally-enforced, nationality-based wage discrimination is massive compared to wage discrimination based on other socially constructed categories irrelevant to intrinsic productivity, such as race or sex/gender. This gain to a mover is roughly double the average GDP per capita of all developing economies in aggregate (P\$4,911 in 2007). Thus it is not surprising that the per-person gains in income or poverty reduction from available public policies or programmatic interventions are tiny fractions of the gains from relaxing the obstacles to the movement of poor people.

## 2 Wage ratios for observably identical workers across the border

We begin by estimating  $R_o$ , the ratio between what a typical worker earns in the United States and what an *observably identical* worker earns in each of 42 developing countries. In the following section we will turn to  $R_e$ , the ratio between US and foreign earnings for an *equivalent* (observably and unobservably identical) worker. Only  $R_e$  is the “place premium”.

### 2.1 Data

As used here, “observably identical” means a person of the same country of birth, same country of education, same level of education, same age, same gender, and same classification of dwelling as “rural” or “urban”. Controlling for these observable traits is made possible by a new, standardized collection of individual level data sets on wage-earners compiled by the World Bank,<sup>3</sup> combined with the US Census Public Use Microdata Sample (PUMS) five percent file.

The unified database describes 2,015,411 individual wage-earners residing in 43 countries close to the year 2000. This comprises 891,158 individuals residing in 42 developing countries, 623,934 individuals born in those same 42 developing countries but residing in the US, and 500,319 individuals born in the US and residing in the US. Each individual record contains the person’s wage in 1999 US dollars at Purchasing Power Parity, country of residence, years of schooling, age, sex, an indicator of urban or rural residence, and indicator variables for the periodicity of the reported wage (weekly, monthly, etc., with monthly as the base group). For those residing in the US, there is additional information on country of birth and year of arrival for the foreign-born. A sampling weight is assigned to each observation indicating the number of individuals in the national population represented and this weight is used in all regressions.

---

<sup>3</sup> The sources for all data are given in the appendix. The basic database is also described in Montenegro and Hirn (2008).

A series of steps brings us from the raw collection of data sets to the estimation sample. First, we remove all self-employed people and unpaid family workers from the data, leaving only wage-earners. This has the advantage of increasing the comparability and accuracy of the earnings measures, but has the disadvantage of eliminating a large portion (though not all) of the informal sector from the sample—especially many agriculturists in the poorest countries. Second, we remove all people aged 14 or less and all people aged 66 or greater. Third, we remove all people reporting zero wage earnings. Fourth, we removed the data from twelve transition countries because many of these countries were undergoing extraordinary instability of prices, wages, and currencies at the time the survey was administered<sup>4</sup>. Fifth, we randomly delete US-born US-residents from the PUMS to reduce the size of that group from about 6.13 million to about half a million, due to binding memory constraints in the microcomputer conducting the statistical analysis, and scale up each person’s sampling weight accordingly. Sixth, we drop Chad from the sample because the sample of US residents in the public-use data does not happen to contain any working-age wage-earners who report being born in Chad. Finally, we drop Honduras from the sample for reasons described below.

The US census data were collected for the year 1999 while the surveys were in the 1990s and early 2000s (only India’s survey was carried out in 1999). We convert each wage estimate in current year local currency to current year US dollars at Purchasing Power Parity using factors from the World Bank (2007) and then deflate these dollar amounts to 1999 PPP US dollars using the PPP factor deflator.<sup>5</sup> To the extent that real wages rose (or fell) relative to the US between 1999 and the year of a country’s survey, the wage ratios for those countries will be slightly underestimated (or overestimated). Converting to PPP also naturally introduces the possibility that errors in any given country’s PPP calculation could affect the results; note, however, that each of the 42 wage ratios we

---

<sup>4</sup> The twelve we remove are: Armenia, Azerbaijan, Bulgaria, Belarus, Croatia, Hungary, Latvia, Moldova, Republic of Macedonia, Russia, Romania, and Slovakia.

<sup>5</sup> After we carried out our analysis the World Bank announced intentions to retroactively adjust the PPP factors we use, but these were unavailable at the time of writing. We note, however, that the most important adjustments foreseen are those to India’s and China’s PPP factors, both of which will tend to lower the PPP dollar-value of non-migrants’ earnings and therefore make the wage ratios reported here tend to underestimate the true ratios. In general, pre-2005 PPP ratios may suffer from a failure to properly control for quality of items priced, “leading to an understatement of price levels in poor countries and to an overstatement of their output and income levels” (Deaton and Heston 2008). This tends to bias our estimated wage ratios downward.



calculate is independent of any data from the 41 other countries. Thus any error in any one country's PPP rate does not propagate to the other estimates.

## 2.2 Method

We compare workers residing in one pair of countries at a time—the US and another country  $j \in J$ —estimating a separate wage regression for each country  $j$ . For example, we can use the estimated coefficients to predict the average wage of a Guatemalan-born, Guatemalan-educated, 35 year-old urban male wage-worker with 9 years of education who resides in the United States, and compare this to the predicted average wage of a 35 year-old urban male wage-worker with 9 years of education in Guatemala. This same analysis is then replicated for each of the 41 other migrant-origin countries. The estimates for each country stand alone and are not influenced by any data quality, conceptual (e.g. similarity of the earnings concept), or empirical difficulties specific to any other country.

The regression specification for each country  $j$  is

$$\ln w_{ij} = \begin{pmatrix} \eta_0 + \boldsymbol{\beta}_0 \mathbf{s}_{ij} + \boldsymbol{\gamma}_0 \mathbf{a}_{ij} + \delta_0 f_{ij} + \zeta_0 r_{ij} \\ \eta_j^r + \boldsymbol{\beta}_j^r \mathbf{s}_{ij} + \boldsymbol{\gamma}_j^r \mathbf{a}_{ij} + \delta_j^r f_{ij} + \zeta_j^r r_{ij} \\ \eta_j^l + \boldsymbol{\beta}_j^l \mathbf{s}_{ij} + \boldsymbol{\gamma}_j^l \mathbf{a}_{ij} + \delta_j^l f_{ij} + \zeta_j^l r_{ij} \\ \eta_j^e + \boldsymbol{\beta}_j^e \mathbf{s}_{ij} + \boldsymbol{\gamma}_j^e \mathbf{a}_{ij} + \delta_j^e f_{ij} + \zeta_j^e r_{ij} \end{pmatrix}' \begin{pmatrix} 1 \\ I_{ij}^r \\ I_{ij}^l \\ I_{ij}^e \end{pmatrix} + \varepsilon_{ij}. \quad (1)$$

where  $w_{ij}$  is the wage of person  $i$  in country  $j$ . The first vector on the right-hand side describes a large number of coefficients and dummy variables reflecting levels of education, age, sex, and rural/urban residence. Starting with the first row,  $\mathbf{s}_{ij}$  is a  $5 \times 1$  vector of dummy variables equal to 1 if the person has completed each of five levels of education,<sup>6</sup> and  $\mathbf{a}_{ij}$  is a  $9 \times 1$  vector of dummies for different age levels.<sup>7</sup> The dummy  $f_{ij}$

<sup>6</sup> Beyond a base group of zero years of schooling, the five categories are 1) 1-4 years, 2) 5-8 years, 3) 9-12 years, 4) 13-16 years, and 5) 17-28 years.

<sup>7</sup> Beyond a base group for age 15-19, the nine age categories are 1) 20-24, 2) 25-29, 3) 30-34, 4) 35-39, 5) 40-44, 6) 45-49, 7) 50-54, 8) 55-59, 9) 60-65 (intentionally includes 65).

indicates female and  $r_{ij}$  indicates residence in a rural area. The  $\eta$ ,  $\delta$ , and  $\zeta$  are coefficients, while the  $\beta$  are  $1 \times 5$  vectors of coefficients and the  $\gamma$  are  $1 \times 9$  vectors of coefficients.

The other rows of the first vector, after dot product with the rightmost column vector, allow all of the estimated coefficients to differ between US-born US-residents, foreign-born US residents who arrived before age 20, foreign-born US-residents who arrived at or after age 20, and foreign residents. The “1” in the first row of that rightmost vector signifies that the base group is US-born residents of the US.  $I_{ij}^r$  ( $r$  for “resident” of country  $j$ ) takes the value 1 if individual  $i$  resides in country  $j$ , or 0 otherwise; these are people born in foreign, residing in foreign.  $I_{ij}^l$  ( $l$  for a “late” arriver) is 1 if individual  $i$  was born in country  $j$ , now resides in the US, and arrived in the US at or above age 20, and 0 otherwise.  $I_{ij}^e$  ( $e$  for “early” arriver) takes the value 1 if individual  $i$  was born in country  $j$ , now resides in the US, and arrived in the US below age 20, and 0 otherwise.<sup>8</sup>

### 2.3 Estimates of $R_o$

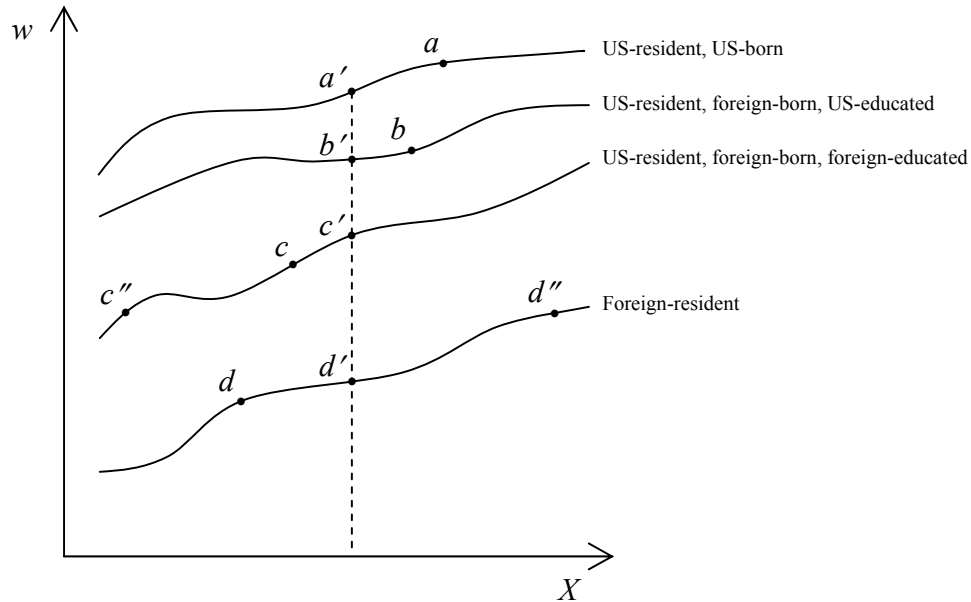
We now present estimated wage ratios based on coefficient estimates from regression (1). Figure 1 provides a schematic visual explanation of the different ratios, and Table 1 presents the estimated ratios. These lead up to our preferred estimate of  $R_o$  in column 6 of Table 1.

In Figure 1, the  $X$  axis shows some observable trait, such as years of education, and the  $w$  axis shows the wage profile associated with that trait—a profile that can take any form. Letting  $a$  represent the vertical height of the point with that label,  $a$  gives the average wage of US-born US-residents,  $b$  is the wage of foreign-born, US-educated US residents,  $c$  is the average wage of foreign-born, foreign-educated US residents, and  $d$  is the average wage of foreign residents.

---

<sup>8</sup> The regressions also include dummy variables for the periodicity of wage reported (daily, weekly, monthly, etc.), suppressed here for clarity.

**Figure 1: Schematic representation of wage ratios**



Each of these has different average wage levels in part because they are at different points on the wage profile in observable trait  $X$ ; people in the US might have a greater number of years of education on average, for example, than people in the foreign country. Thus column 1 of Table 1 gives the ratio  $a/d$  in Figure 1; column 2 gives  $b/d$ ; and column 3 gives  $c/d$ . These first three columns do not control for any observable traits besides country of birth and age of arrival (that is, the coefficients  $\beta$ ,  $\alpha$ ,  $\delta$ , and  $\zeta$  are constrained to zero, and only the coefficients  $\eta$  are estimated).

In other words, column 1 shows the ratio between the average monthly wage of a US-born, US-resident worker and a foreign-resident worker, without controlling for any observable traits. In column 2, the numerator is the average wage of a foreign-born, US-resident worker, without controlling for any other traits. In column 3, the numerator is the average wage of a foreign-born, US-resident worker who arrived at or after age 20, again, without controlling for other traits.<sup>9</sup>

<sup>9</sup> The estimates in column 2 are closely related to those of Hendricks (2002) who uses these to adjust cross-national estimates of human capital for growth accounting.

**Table 1: Estimates of wage ratios for observably identical workers ( $R_o$ )**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Numerator is:	X			X				
<i>US-born</i>		X			X			
<i>Foreign-born</i>			X			X		X
<i>Foreign-born &amp; educ.</i>								
Controls				X	X	X		X*
							Conf. Interval for column 6	
Yemen	14.06	11.18	12.20	18.02	16.64	<b>15.45</b>	(5.58, 42.79)	11.43
Nigeria	13.45	15.52	17.34	21.93	16.21	<b>14.85</b>	(10.28, 21.46)	7.79
Egypt	11.98	16.00	17.36	15.89	11.97	<b>11.92</b>	(7.43, 19.12)	11.93
Haiti	23.50	20.48	21.40	16.06	10.60	<b>10.31</b>	(7.79, 13.67)	4.19
Cambodia	9.16	8.49	9.12	8.45	8.11	<b>7.45</b>	(4.66, 11.91)	6.40
Sierra Leone	8.35	8.06	9.29	10.92	7.61	<b>7.43</b>	(0.2, 280.18)	3.70
Ghana	9.37	10.41	11.53	9.68	7.28	<b>7.12</b>	(1.27, 40.04)	4.22
Indonesia	9.52	9.38	10.92	8.89	6.77	<b>6.72</b>	(3.39, 13.32)	3.17
Pakistan	12.58	13.65	15.32	8.40	6.88	<b>6.57</b>	(4.88, 8.85)	2.95
Venezuela	8.89	8.94	8.97	8.65	8.08	<b>6.57</b>	(4.18, 10.30)	3.69
Cameroon	10.06	10.36	12.11	8.44	5.81	<b>6.53</b>	(2.09, 20.44)	7.38
Vietnam	10.29	9.97	9.92	8.88	7.39	<b>6.49</b>	(5.56, 7.56)	3.92
India	10.88	16.33	18.49	7.37	6.69	<b>6.25</b>	(5.28, 7.39)	2.96
Jordan	6.21	7.02	7.24	7.68	6.25	<b>5.65</b>	(2.77, 11.50)	3.98
Ecuador	7.29	6.17	6.01	7.89	5.75	<b>5.16</b>	(3.99, 6.67)	3.26
Bolivia	5.78	5.51	5.75	7.09	5.58	<b>5.03</b>	(2.76, 9.18)	3.34
Sri Lanka	9.90	11.95	12.76	7.32	5.22	<b>4.95</b>	(2.34, 10.49)	1.26
Nepal	13.47	10.61	12.14	9.08	5.30	<b>4.85</b>	(1.45, 16.19)	4.37
Bangladesh	6.89	6.23	6.90	6.56	4.62	<b>4.60</b>	(2.96, 7.14)	2.19
Uganda	7.71	10.01	10.52	6.00	3.61	<b>4.38</b>	(1.31, 14.64)	2.30
Ethiopia	13.07	12.08	13.01	6.16	4.46	<b>4.35</b>	(2.81, 6.73)	2.40
Guyana	4.81	5.29	5.77	4.61	4.06	<b>3.87</b>	(2.06, 7.24)	1.39
Philippines	6.19	7.38	8.27	5.63	4.44	<b>3.82</b>	(3.36, 4.35)	1.42
Peru	4.43	4.11	4.29	5.01	4.08	<b>3.79</b>	(2.96, 4.85)	1.60
Brazil	5.03	4.85	5.29	4.31	3.90	<b>3.76</b>	(2.88, 4.92)	1.66
Jamaica	3.20	3.45	3.80	4.63	3.88	<b>3.63</b>	(0.67, 19.79)	1.55
Chile	3.09	3.39	3.64	4.28	3.70	<b>3.53</b>	(2.17, 5.76)	1.60
Nicaragua	4.96	3.94	4.12	5.10	3.91	<b>3.52</b>	(2.61, 4.75)	1.42
Panama	3.86	4.18	4.14	4.46	3.94	<b>3.36</b>	(2.06, 5.49)	1.54
Uruguay	3.22	3.66	3.72	2.72	3.26	<b>3.10</b>	(1.28, 7.50)	1.90
Guatemala	6.25	4.65	4.78	4.28	3.07	<b>2.94</b>	(2.39, 3.61)	1.73
Colombia	4.42	3.87	3.87	4.14	3.11	<b>2.88</b>	(2.40, 3.46)	1.65
Paraguay	3.20	3.21	3.48	2.74	2.65	<b>2.78</b>	(0.71, 10.93)	1.10
South Africa	2.83	4.01	5.18	2.34	3.04	<b>2.75</b>	(1.45, 5.21)	0.65
Turkey	3.15	3.48	3.86	3.55	3.03	<b>2.68</b>	(1.52, 4.74)	1.46
Argentina	2.40	2.88	2.92	3.05	2.90	<b>2.54</b>	(1.60, 4.04)	1.37
Mexico	3.82	2.72	2.68	3.83	2.78	<b>2.53</b>	(2.42, 2.65)	1.31
Belize	3.58	3.49	3.65	2.59	2.38	<b>2.43</b>	(0.02, 243.65)	1.16
Thailand	4.79	3.80	4.86	3.22	2.65	<b>2.17</b>	(1.29, 3.64)	1.04
Costa Rica	2.85	2.62	2.52	2.92	2.25	<b>2.07</b>	(1.22, 3.53)	1.24
Morocco	3.48	4.06	3.81	3.02	2.37	<b>2.00</b>	(1.06, 3.74)	0.62
Dominican Rep.	3.32	2.44	2.52	3.02	2.04	<b>1.99</b>	(1.66, 2.39)	1.30
<i>Median</i>	<i>6.20</i>	<i>5.84</i>	<i>5.89</i>	<i>5.81</i>	<i>4.26</i>	<i>4.11</i>		<i>1.81</i>
<i>Mean</i>	<i>7.27</i>	<i>7.38</i>	<i>7.99</i>	<i>6.88</i>	<i>5.43</i>	<i>5.11</i>		<i>2.99</i>

Sorted in descending order by column 6. Columns 1-3 give the ratio average wage of a worker residing in the US to the average wage of a worker residing in each foreign country, without controlling for observable traits besides country of birth and age of arrival. Columns 4-6 give the predicted ratio between the average wage of a US-resident 35 year-old male urban worker born in each country with 9 years of education acquired in each country, to the average wage of an *observably identical* worker residing in each origin country. Column 7 gives a 95% confidence interval for column 6. \*Column 8 is identical to column 6 except the numerator contains the predicted wage for a person in the US that has completed only primary education, while the denominator contains the predicted wage for a person in the foreign country who has completed four years of tertiary education.

The next three columns of Table 1 control for observable traits (education, age, sex, and rural/urban) and graphically are various ways of “drilling down” through the wage profiles to compare the predicted average wages of persons at the same point in observed characteristics. The specification of regression (1) allows different wage profiles for each of the four groups (so the curves in Figure 1 are not forced to be parallel). Column 4 of Table 1 gives the ratio  $a'/d'$  in Figure 1, column 5 gives the ratio  $b'/d'$ , and column 6 gives the ratio  $c'/d'$ , or  $R_o$ .

Ratios in the remaining columns control for education, age, gender, and rural/urban residence. They are based on empirical estimates of the parameters  $\beta$ ,  $\alpha$ ,  $\delta$ ,  $\zeta$ , and  $\eta$ , and give predicted average wage for a 35 year-old urban male with 9 years of education. In column 4 the numerator is once again US-born US-residents, and in column 5 it is foreign-born US-residents. In column 6 the numerator represents workers born in each country of origin and (likely) educated there, having arrived at or after age 20. These ratios, in boldface, are the estimates of  $R_o$ —the ratio of predicted wages for observably identical workers across the US border. Column 7 gives a 95% confidence interval for the point estimates in column 6, based on a simple  $F$ -test of coefficient restrictions in regression (1). The raw coefficient estimates used to calculate Table 1 are given in Appendix Table A1.

The median estimated  $R_o$  in column 6 is around 4.0, corresponding roughly to Ethiopia, Peru, or Guyana. The highest estimated  $R_o$  is for Yemen at 15.45 (earning \$1,940 per month in the US versus \$126 per month in Yemen), while the lowest is for the Dominican Republic (earning \$1,491 per month in the US versus \$749 in the Dominican Republic).<sup>10</sup> The highest absolute difference in annual wage earnings is \$21,722 (Yemen), the smallest is \$8,912 (Dominican Republic). The mean and median annual absolute differences are both just over \$15,400. Comparing columns 1 and 6 reveals that observable individual traits typically explain about one third of international differences

---

<sup>10</sup> The wage premia tend to be modestly lower at higher levels of education (although this is in ratios; in absolute terms the gap grows). This can be attributed mechanically to the fact that the partial association of wages in the US labor market and schooling acquired abroad (median 6.1% increase in wages per year of schooling) is typically substantially lower than the association of US wages and US schooling (median 12.3%) or the association of foreign wages and foreign schooling (median 8.2%).

in wages, as the median raw wage ratio is 6.2 and the median of the ratio for observably identical workers is 4.1.<sup>11</sup>

The enormous size of the ratios  $R_o$ , compared to wage differences created by differences in other wage determinants such as education, is underscored by column 8 of the table. This is identical to column 6 except for one change: It compares the average predicted wage of a foreign-born, foreign-educated, 35 year-old urban male in the US *who has completed only primary education* to the average predicted wage of a 35 year-old urban male in the foreign country *who has completed four years of tertiary education* (interpreting  $X$  in Figure 1 as education, column 8 thus shows the ratio  $c/d$ ). For example, an average Indian worker with six years of Indian education earns about triple the wages working in the United States, adjusted for purchasing power, as a person with 16 years of education earns in India.

### 2.3 Robustness of the estimated $R_o$

As with any empirical exercise, we make a number of assumptions. Here we discuss several of these assumptions and their possible effects on the magnitude of the results.

*Exchange rates:* By using PPP exchange rates we are implicitly assuming that all consumption of movers occurs in the US, which substantially understates the gains to overall earnings for migrant families, in two ways. First, this ignores remittances. If a worker is in one country with nuclear family members in another, and if we assume a unitary household utility function, then *household* consumption should be deflated in the location where consumption occurs. This suggests at the least that all remittances should enter the analysis at sending country prices (official exchange rates), not PPP. Second, migrants, and especially temporary workers, should optimally have very high savings rates. A simple model of inter-temporal consumption smoothing would suggest that if a

---

<sup>11</sup> Median:  $1 - (4.11/6.20) = 33.7\%$ . Mean:  $1 - (5.11/7.27) = 29.6\%$ . Milanovic (2008) shows that country fixed effects explain roughly 60 percent of all income inequality across individuals in the world, but this includes inequality due to differential access to capital and different levels of human capital. In contrast, our results are specific to labor income for workers with the same characteristics.

worker had access to a much higher wage rate for an explicitly temporary period they should optimally smooth these windfall gains over his or her lifetime. Alternatively, temporary migration is often modeled as driven by “target savers” who accumulate savings for a specific purpose (e.g. a house, business, car, wedding/marriage), consumption that, again, would occur in their country of origin not in the US. Much, perhaps most consumption of the US earnings of temporary migrants would be in their own country, not the US.

Table 2 explores the sensitivity of the  $R_o$  estimates to the choice of exchange rate. The leftmost column uses PPP exchange rates and reproduces the estimates in column 6 of Table 1. This is equivalent to the assumption that none of the increase in earnings of movers is spent at origin-country prices. The rightmost column shows what  $R_o$  would be if official exchange rates are used, equivalent to the assumption that all of the increase in earnings is spent in the origin country. The two intermediate columns use two different weighted-average exchange rates. Column 2 assumes that roughly 20% of migrants’ income is spent in the origin country, a conservative estimate for Mexicans in the US.<sup>12</sup> Column 3 assumes that 60% is spent at the origin, in line with estimates for male overseas Filipino contract workers.<sup>13</sup>

The median estimate of  $R_o$  rises from about 4 using PPP to about 5 when 20% of income is spent in the origin country, to above 7 when 60% of income is spent, and to 14 when using official exchange rates. Even the three smallest wage ratio countries at PPP are above 3 at 60 percent and above 5 at official exchange rates. The estimates of  $R_o$  in Column 6 of Table 1 are conservative in potentially substantially *understating* the “real” wage differentials based on the relevant consumption prices of movers.

---

<sup>12</sup> Amuedo-Dorantes, Bansak, and Pozo (2005, Table 1A) find that Mexican migrant household heads in the United states remit 27.9% of monthly income to Mexico, a figure that includes non-remitters and does not include repatriated savings.

<sup>13</sup> Semyonov and Gorodzeisky (2005, Table 1) find that male overseas Filipino workers remit 60.3% of monthly income to the Philippines, while females remit 45.0%.

**Table 2: Sensitivity to choice of exchange rate**

Estimates of  $R_o$  using PPP exchange rates (leftmost column), official exchange rates (rightmost column), or weighted averages of the two (intermediate columns)

Representative countries	Placement of countries in the distribution baseline $R_o$	% of income consumed in country of origin			
		0% (PPP- Column 6 of Table 1)	20%	60%	100% (Official exchange rate)
Yemen	} Three highest	15.45	16.05	17.40	19.00
Nigeria		14.85	16.79	22.71	35.08
Egypt		11.92	13.49	18.33	28.58
Cameroon	} Around 75 <sup>th</sup> percentile	6.53	7.56	11.06	20.57
Vietnam		6.49	7.75	12.72	35.42
India		6.25	7.44	12.01	31.15
Guyana	} Around Median	3.87	4.58	7.30	17.88
Philippines		3.82	4.53	7.16	17.09
Perú		3.79	4.29	5.83	9.08
Colombia	} Around 25 <sup>th</sup> percentile	2.88	3.33	4.81	8.67
Paraguay		2.78	3.26	4.94	10.21
South Africa		2.75	3.17	4.55	8.08
Costa Rica	} Three Lowest	2.07	2.31	3.00	4.27
Morocco		2.00	2.28	3.19	5.33
Dominican Rep.		1.99	2.28	3.24	5.55
<i>Median of all 42 countries</i>		<i>4.11</i>	<i>4.92</i>	<i>7.23</i>	<i>13.90</i>

*Reliability and comparability of reported earnings:* Research comparing multiple sources of income data at the individual level suggest that self-reported income is an unbiased estimator of true income, both in rich countries (Bound and Krueger (1991)) and in poor countries (Akee (2007a)). There is less certainty about comparability. Wage data for the US reflect total earnings from all jobs, whereas wage data for the 42 developing countries in our sample reflect wages from the respondent's principal occupation. For the vast majority of formal-sector wage earners in the sample we nevertheless expect wage earnings from the principal occupation to closely reflect total wage earnings.

Furthermore, wage data for the United States reflect gross earnings before taxes, and we expect that most people responding to a general question about their wages or earnings would have provided gross wages on most of the country surveys, but for a handful of



countries it may be that the responses reflect after-tax wages.<sup>14</sup> If respondents provided net-of-tax instead of gross wages this would result in some upward bias to our estimated  $R_o$ . This bias will be small, however, if it is present at all. Formal-sector income taxes are on the order of 5% in most developing countries (Easterly and Rebelo (1993)). For the median ratio of 3.92, for example, a 5% underestimation of the denominator means that the corrected ratio is 3.73.

Reported wages in the US census do not include non-wage benefits, which are likely to be a larger fraction of total compensation in the US than in many of the countries examined here. Again most of these considerations of comparability would tend to make  $R_o$  *underestimate* the cross-border ratio of total compensation.

*Regression specification:* Heckman, Lochner, and Todd (2006) question the validity of assumptions underlying the traditional Mincer functional form, which helps motivate our choice of the much more flexible specification in (1). We also conducted the same analysis using (i) the traditional Mincer specification and (ii) augmented Mincer specification with square and cubic terms and interactions allowing a flexible approximation of the functional form of more complex education-wage and age-wage relationships. All the variations in functional form we experimented with gave almost identical overall results for the one comparison group we chose<sup>15</sup> and hence are omitted.

The reported estimates ratios  $R_o$  are just factual summary statistics about wage data, the ratios of the predicted conditional means of two wages of two different groups—people who are the same in the characteristics in the two samples—on opposite sides of the US border. These wage ratio estimates are almost certainly conservative and are robust to the

---

<sup>14</sup> In a small number of the countries (such as Yemen) the survey explicitly requests after-tax earnings, and in a few of the others (such as Chile) custom may dictate that formal sector “wages” refer to after-tax earnings unless otherwise specified. The text of the wage question from each survey is in the Appendix.

<sup>15</sup> Which is not to say: “functional form doesn’t make a difference” in estimating wage profiles. Many of the functional form assumptions affect the slopes of the wage profiles in Figure 1, but if we are “drilling down” at a single point near the middle of the education distribution (as opposed to say, comparing wage differentials across countries) one can imagine a good deal of robustness even if functional form does matter for other questions.

functional form used to estimate the wage profiles used in computing the conditional means.

These facts themselves have never before been recorded on such a wide scale. As with most empirical work in economics, all the theory, and controversy, comes in interpreting these facts. The wage data represent the outcomes of the workings of spatially separated labor markets, which themselves are the result of the choices of employers in each of those markets, choices of workers in each of those markets, and choices of workers to (attempt to) move across markets, and all of these choices are constrained by institutions and policies—including policies about crossing the border.

### **3 Assessing wage ratios for fully equivalent workers**

The preceding estimates of  $R_o$  could be biased estimates of what we term the “place premium” the cross-border ratio of wages earned by two people of equal intrinsic productivity. In particular, we are interested in measuring  $R_e$  for the marginal person who would cross the border if policy barriers were incrementally relaxed. As we will show,  $R_o$  overestimates  $R_e$  principally to the extent that migrants are positively selected on unobservable determinants of wage. If, through choices of movers, employers or policy, the workers in the US would have had above average earnings in their home-country labor market because of unobserved wage determinants not included in our regressions, then the ratio  $R_o$  overestimates  $R_e$ .

In this section, we first describe the form of this bias. We then present new evidence from a variety of sources on the degree of this bias. We proceed to triangulate our findings with the existing evidence. Finally, we calculate  $R_e$  using the actual distributions of residuals under different assumptions about the degree of selectivity suggested by the preceding evidence. While no one piece of this evidence is definitive, the preponderance of the evidence suggests that, among low to moderate skilled workers, the selection is

positive, but not very strongly so.<sup>16</sup> The extent to which the wage ratios of observably identical workers overstates the wage ratios of equal-productivity marginal movers is generally modest, with  $R_o/R_e$  falling in the range 1.0 to 1.5 in all cases examined.

### 3.1 *Migrant selection and estimating $R_e$*

Suppose that each potential migrant has an idiosyncratic wage at the origin (home country  $h$ ), wage at the destination (country  $d$ ), and cost of moving, broadly considered, which includes the obstacles and costs created by policy. The marginal migrant will be one for whom the wage gain to movement just equals the moving cost:

$$\ln \bar{w}_d + \ln \tilde{\mu}_d^i - (\ln \bar{w}_h + \ln \tilde{\mu}_h^i) = \ln \pi^i . \quad (2)$$

Here,  $\bar{w}_d$  is the average wage at the destination earned by a person from the origin country *for a given set of observable traits*, and  $\bar{w}_h$  is the average wage earned by an observably identical person in the home country of origin.  $\tilde{\mu}_d^i$  is the unobservable difference between the wage that will be earned at the destination by a marginal observably identical migrant  $i$  and the average previous observably identical migrant.  $\tilde{\mu}_h^i$  is the unobservable difference between the wage that this same person would earn in the home country of origin and the average earnings of an observably identical person at the origin. Finally,  $\pi^i$  is that person's cost of moving.

After taking expectations of both sides of (2), assume that the marginal migrant can expect roughly the same wage outcomes as previous observably identical migrants:  $E[\ln \tilde{\mu}_d^i] \approx 0$ . Taking the exponent of both sides, letting  $\mu_h \equiv \exp(E[\ln \tilde{\mu}_h^i])$ , and letting  $R_e \equiv \exp(E[\pi^i])$ , we have

---

<sup>16</sup> This is not a general claim about “selectivity” in mobility decisions as we are not examining college graduates, much less the highly skilled “superstar” movers such as economics professors and our data does not distinguish between legal and undocumented workers in the US. It is possible that selectivity is a much larger issue for legally admitted and/or higher-skill workers, who are a focus of the “skill price” approach in Rosenzweig (2006).

$$R_o \equiv \frac{\bar{w}_d}{\bar{w}_h} = \mu_h R_e. \quad (3)$$

The left-hand side of (3) is the destination-to-origin wage ratio for an observably identical worker—what we call  $R_o$ . On the right-hand side,  $R_e$  is the “place premium”, the wage ratio for an equivalent intrinsic productivity worker across the border. The term  $\mu_h$  reflects selection of the marginal migrant from the distribution of unobservable determinants of earnings in the migrant’s home country. If he or she comes from above the conditional mean of the unobservable determinants of earnings, then  $\mu_h > 1$  and the ratio  $R_o = \bar{w}_d / \bar{w}_h$  overestimates the place premium  $R_e$ .

As a heuristic description, we are not estimating the “average treatment effect” or the wage gain if the “typical” Peruvian worker were involuntarily moved to the US labor market. In this case one would worry that the wages of the actual movers would overstate the gain to the average Peruvian worker, if moved, because movers had self-selected because their wages in the US would be high—e.g. they had language skills or relatives who could locate jobs—so that the average wage of the existing movers would be in the tail of the distribution of unobserved determinants of wages in the US. But we are interested in the marginal *voluntary* mover if the distribution of moving costs were incrementally proportionally reduced for all potential movers. In this case we assume that the difference of the marginal mover and existing movers in the US labor market is small. But it is still the case that both the existing movers and the marginal mover might have had much higher wages had they remained in the *home* market and hence comparing the marginal worker by comparing the average of existing (late arriving) movers to non-movers gets badly wrong the wages the mover *would have* earned had they remained. So even though we know the wages of movers in the US and the wages of non-movers in their home country, the question is: What would have been the wages of the movers had they not moved?

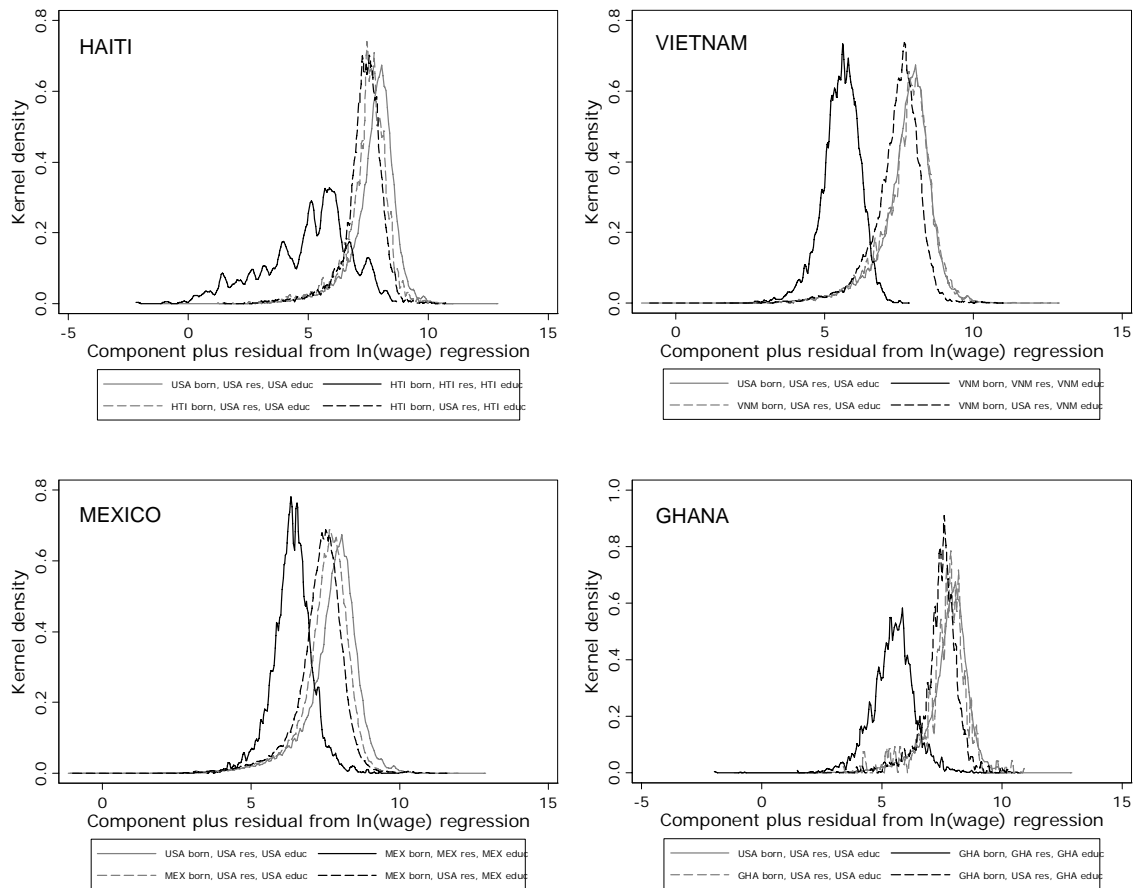
### 3.2 *Sensitivity to assumptions about selection*

What degree of selection would be necessary to produce substantial differences between  $R_o$  and  $R_e$ ? Figure 2 shows sample kernel density plots of the predicted distributions of formal sector wages for 35 year-old urban males with 9 years of education for (i) US-born US residents, (ii) foreign-born US-resident “early arrivers” (before age 20), (iii) foreign-born US-resident “late arrivers” (at or after age 20, thus almost all educated abroad), and (iv) foreign residents, for four representative countries: Mexico, Vietnam, Ghana, and Haiti.

$R_o$ , as estimated in column 6 of table 1, is essentially the ratio of the means of the last two of these groups (“late arrivers” versus “non-movers”). Hence Figure 2 shows the distribution around the vertical slices through the wage profiles in Figure 1 (as the lines in that figure are the regression function) and illustrates that the distribution of “late arrivers” and “non-movers” are spaced far apart—in some cases mostly non-overlapping. Even comparing the 50<sup>th</sup> percentile of the “late arriver” distribution to the upper percentiles of the “non-mover” distribution, consistent with strong positive selection, would still produce very large wage ratios.

What should be our prior about the degree of selectivity of existing movers from their home country *residual* distribution of wages from the countries under consideration? The observed degree of selectivity is the result of a variety of factors in the movement decisions of individuals. That is, among the low to moderate skill workers the marginal migrant has not been purposefully “selected” for entry into the US based on characteristics that are likely positively correlated with the unobservable component of wages in the home country. This is likely true of at least some people reading this paper, who obtained visas or citizenship based on “extraordinary” or “exceptional” ability, or of H1-B visa holders who are chosen based on demand from employers, but is less true of the typical high school or less educated Mexican or Bolivian or Vietnamese worker in the sample.

**Figure 2: Kernel densities of the unexplained component of wages**



The observed selectivity in equilibrium of actual movers is based on a combination of individual specific characteristics which determine the propensity to move, such as variations in the policy based constraints to movement, variations in the ease of evading those policies and entering the country and worked as an undocumented worker, networks or connections of friends and relatives in the US that lower the job search and/or psychic costs of moving, the utility loss to being abroad, and so on. Some of these factors driving the movers' decisions might be correlated with home country wages, some not. All else equal, lower home country wages would lead to a higher propensity to move. There is nothing about the theory or practice of labor mobility across the US border that guarantees movers will be strongly positively selected, especially in the

relevant sense here of moderate skill workers coming from high percentiles of the distribution of wages conditional on observed characteristics.<sup>17</sup>

Hence we believe there is scant existing theoretical or empirical basis to form priors that selection on unobservables from the countries and for the workers under consideration is negative, neutral, or positive, much less of its magnitude. It is not even clear what the relevant stance of “methodological skepticism” about econometrics and identification would be in this case. One might frame the question such that the null is “no selection” and hence in the absence of ironclad evidence to the contrary the default is that the observed estimates of  $R_o$  should not be adjusted for selection at all. On the other hand, one could frame the null as  $R_e = 1$  for all countries (or the “no policy barriers” level)—ignoring evidence of gains to migration and without any theoretical rationale—and then demand ironclad evidence to be swayed from this null. But reconciling  $R_e = 1$  for all countries with the observed  $R_o$  from Table 1 requires asserting a complex set of priors, such that the magnitude of selection for each country is just exactly what is required to reconcile the null and the observed  $R_o$ . That hardly seems methodologically defensible.

We move ahead by providing several distinct ways of estimating the degree of positive selection, from different methods and countries. None of this evidence allows a definitive, ironclad point estimate of  $R_e$  for any country. But it is sufficient to allow estimation of a reasonable prior for  $R_e$  in several of the countries we examine, as well as to establish reasonable priors for the degree of selection and the magnitude of  $R_e$  for typical countries.

### 3.3 *New evidence on selection: Panel data*

One direct approach to estimating the point of the home country distribution of wages from which migrants come is to look at migrants’ wages before they move and compare

---

<sup>17</sup> In fact, a common objection to current US immigration policy is that it has been based far *too little* on productivity related characteristics and far too much on characteristics unrelated to productivity, such as having a relative in the US which has led to a *declining* “quality” of allowed migrant in terms of observables (e.g. Borjas 1987), on top of which the net result of enforcement of existing policy is that in many instances the marginal migrant is likely in the US as entirely undocumented or not in compliance with their original visa entry status.

the residual distribution of wages of those who will later be migrants to those who do not migrate. Existing, publicly-available data allow such an exercise in the Philippines, Mexico, and South Africa.

The Labor Force Survey of the Philippines allows analysis of this kind. The nationally-representative, rotating panel design of quarterly surveys conducted between January 2001 and April 2003 allow construction of seven separate panels of several thousand households lasting nine months each. Wages and basic traits of each worker are collected at the beginning of the panel, and it can be determined whether each worker left the Philippines by nine months later.<sup>18</sup> In these data, there are observed wages for 64,172 different workers at the start of all the panels taken together aged 15 to 65 with an observed wage. 277 of these people had left the Philippines by the time the household was visited again nine months later. We estimate a wage profile as close as possible to regression (1), given the constraints of the data,<sup>19</sup> and calculate the residual.

Figure 3 shows a kernel density plot of the residual in initial wages comparing subsequent movers to non-movers. The mean wage residual of the movers lies at the 54<sup>th</sup> percentile of the distribution of unobserved determinants of earnings for non-mover. To the degree that this is representative of selectivity of Filipinos departing for the United States, this suggests that  $\mu_h$  in equation (3) is 1.08 (95% confidence interval (1.02,

---

<sup>18</sup> Construction of the panel requires a nontrivial matching procedure because the survey does not retain strictly unique household identification numbers over time. Households that are retained in the rotating sample between two periods do retain the same identification number, but new households rotated into the sample are given identification numbers that, in a small number of cases, could be recycled identification numbers of households that could not be contacted and were dropped from the panel (between 6% and 8% of the households per wave). For this reason, we match households across time on the household identification number plus a code number reflecting the demographic composition of the household. This allows matching of households to a high degree of confidence. Within the matched households, subsequent migrants are identified by comparing wage-earners' age and gender at the start of the panel to migrant household members' age and gender at the end of the panel. While it is quite possible for this to generate ambiguity about which of (say) two people in the same household with identical age and gender became the emigrant, this did not happen to occur in the 277 cases of subsequent migrants with initially observed wages in this sample. Despite the minor limitations of the data, therefore, we are very confident of high-quality matching. And even if matching were poor, it would be unclear why false matches would tend to generate bias in the results through any correlation with the degree of migrant selection.

<sup>19</sup> The age dummies are identical to those used in regression (1). The education dummies are in seven categories: 1) no schooling (base group), 2) primary school attained, 3) primary school completed, 4) secondary school attained, 5) secondary school completed, 6) tertiary school attained, 7) tertiary school completed. The regression and the kernel density plot are weighted by sampling weights.



1.15)). Since  $R_o$  for the Philippines is 3.8, using this  $\mu_h$  would imply that for the Philippines  $R_e \approx 3.5$ . (The median mover comes from the 58<sup>th</sup> percentile of non-movers.)

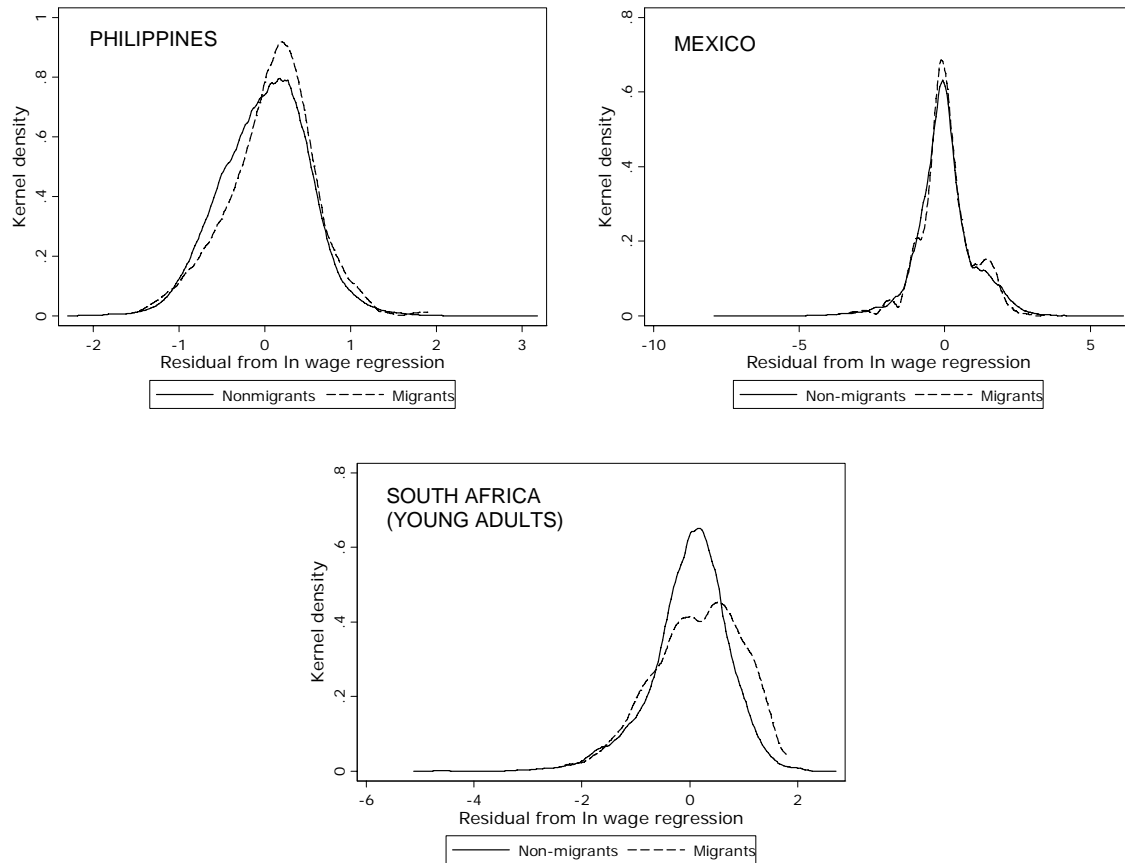
The structure of the Mexican Labor Force Survey (Encuesta Nacional de Ocupación y Empleo) since 2005 allows a nearly identical analysis. Each household that enters the panel is visited five times, allowing observation of workers over a full 12 months. As long as the household can be recontacted in each wave, it can be determined whether or not any household members left the country during the period of observation. The Mexico data from 2005 to 2008 thus contain wage information for 274,955 workers, 569 of which are known to have left Mexico by 12 months later. Again we regress the natural logarithm of wages on the same set of observable traits closely analogous to equation (1)<sup>20</sup> and calculate residuals.

Figure 3 compares kernel density plots for residual wages earned by non-emigrants to those earned by subsequent emigrants. These nationally-representative data reveal that the average emigrant comes from the 56<sup>th</sup> percentile of residual wages, suggesting that  $R_o/R_e = 1.03$  (with a 95% confidence interval of (0.96, 1.12)), so that  $R_e \approx 2.46$ . (The median emigrant comes from the 50<sup>th</sup> percentile of non-migrants.) This calculation for Mexico is important but should not be seen as representative of selection processes for other countries that are more distant, where language barriers are more important, and where diasporas are smaller, all of which might affect the degree of selection.

---

<sup>20</sup> The only difference is that the categorical education dummies represent 1) no schooling (base group), 2) less than primary, 3) primary, 4) secondary, 5) *preparatorio* or *bachillerato*, 6) normal school (teachers' college) degree, 7) technical degree, 8) professional degree, 9) master's degree, and 10) doctorate.

**Figure 3: Selection on unobservable wage correlates in household panel data**



Corroborative but more limited panel data evidence comes from South Africa, where the Cape Area Panel Study gathered wage and international migration data on 2,864 young adults (aged 14-22) in Cape Town, South Africa followed from 2002 to 2006. These data have the advantage that each person was followed for a full four years, but the disadvantage that they cover only young adults. Of those who began the study in 2002, 61 individuals had both left South Africa by 2006 and reported a wage in South Africa prior to departure. We calculate the residual log wage controlling for age, education, and

sex,<sup>21</sup> but do not control for race, and compare this residual for those who did not leave South Africa to the residual for those who did (Figure 3).

In these Cape Town data, the average migrant comes from the 55<sup>th</sup> percentile of the distribution of unobserved determinants of wages among non-migrants.<sup>22</sup> This finding is notable because South Africa is a case where selection of emigrants on wage-related traits unobserved in our earlier regressions (especially race) might be presumed very large. Among young adults, positive selection on unobservables is only moderate, and while the sample is not representative of workers of all ages, it is unclear why selection of somewhat older workers would differ dramatically from that of young adults. If it does not, our estimate of  $R_o = 2.75$  for South Africa likely overstates  $R_e$  by a factor of about 1.20 (95% confidence interval (1.01, 1.41)), so that  $R_e$  would be roughly 2.3.<sup>23</sup> (The median migrant comes from the 60<sup>th</sup> percentile of non-migrants.)

One important drawback of this origin country-based household survey approach is that the data inherently omit information on the emigration of workers whose entire household emigrated and therefore was dropped from the home country-based panel of households. It is not obvious, however, why selection patterns for whole-household emigrants would differ radically from those for partial-household emigrants. Thus the degree of selection observed for partial-household emigrants is informative to shape our priors about the degree of selection among all emigrants.

---

<sup>21</sup> We include a full set of dummies for each age at which the wage was observed (14-26), one dummy for each year of schooling, and a dummy for sex. We omit a rural/urban dummy as that is already held constant; the sampling universe is Cape Town. The regression and kernel density plot are weighted by the “young adult weight”, which adjusts for the sample design, household nonresponse, and young adult nonresponse. The regression also includes a set of dummies for the year in which the reported wage was earned (to control for inflation).

<sup>22</sup> The mean residual log pre-migration wage among migrants weighted by sampling weight is 0.145. The 55<sup>th</sup> percentile of the wage residual, weighted by the sampling weight, is 0.142. While the study does not report the destination of the migrants, 79% of the migrants are described as “white” even though whites are just 13% of the non-migrant sample. This strongly suggests that the destination of most emigrants was a rich country such as the United Kingdom, Canada, Australia, or the United States. This is verified by other research (e.g. Shaw 2007, p. 8) and by the government of South Africa (e.g. in Statistics South Africa [2005], *Documented Migration 2003*, Report No. 03-51-03. Pretoria: Statistics South Africa).

<sup>23</sup> When the same regression is re-run with a dummy indicating which workers subsequently became migrants, the coefficient on this dummy is 0.179 (with a standard error of 0.085), and  $e^{0.179} = 1.196$ .

A second drawback is that there is no wage observation for people who were not employed prior to departure. To the extent that people who emigrate are any *less* likely to be employed prior to emigrating, this not a major concern if the purpose of the exercise is to place an upper bound on the degree of positive selection: including them would only make selection more negative. To the extent that emigrants are greatly *more* likely to be employed prior to emigrating than nonmigrants, the degree of positive selection would be understated. But we see no evidence that this is a major generalized phenomenon in the Philippines or Mexico.

### 3.4 *New evidence on selection: Wage histories of movers*

Panel wage data on international migrants are scarce. Somewhat more common are data on the wage histories of international migrants. The Mexican Migration Project and Latin American Migration Project have collected work histories of several thousand migrants from a variety of Latin American countries. These include a datum for each migrant household head or spouse on that person's last wage in the country of origin. Because the surveys collect identical data on large numbers of nonmigrant households as well, this allows comparison of migrants' last observed wage in the home country to that of nonmigrants, holding observable traits constant.<sup>24</sup>

---

<sup>24</sup> The raw data do not indicate a year in which the last home-country wage was earned, but the labor history of each individual allows determination of the last year in which each respondent reports working in the home country. Based on this we deflate each last home-country wage to constant currency units for the year 2000, using the Consumer Price Index for each country from the World Bank's *World Development Indicators 2008*. We use only individuals with wage observations after 1980 (or after 1991 in the case of Nicaragua, due to a preceding hyperinflation). While the original wages are reported in several different periodicities (hourly, weekly, etc.) we convert to monthly wages by assuming an 8-hour workday and 6-day workweek. For 15 individuals we impute the missing periodicity based on the magnitude of reported wage. For Peru we omit migrants who went to countries other than the US, Mexico, Spain, or Costa Rica.

**Table 3: Estimates of  $R_e$  for in Latin America from wage histories**

	Mexico	Guatemala	Nicaragua	Costa Rica	Dom. Rep.	Haiti	Peru
Number of obs.	8,102	552	1,947	1,535	1,046	405	780
<i>of which migrants surveyed abroad</i>	<i>936</i>	<i>34</i>	<i>173</i>	<i>97</i>	<i>99</i>	<i>20</i>	<i>18</i>
	<i>140</i>	<i>0</i>	<i>81</i>	<i>21</i>	<i>50</i>	<i>0</i>	<i>0</i>
Typical migrant percentile in distribution of non-migrants' unobserved component of wages							
Mean migrant:	54	47	53	61	51	60	69
Median migrant:	49	54	50	55	50	56	70
<b><math>R_o</math></b>	<b>2.53</b>	<b>2.94</b>	<b>3.52</b>	<b>2.07</b>	<b>1.99</b>	<b>10.31</b>	<b>3.79</b>
<b><math>R_e</math></b>	<b>2.35</b>	<b>3.08</b>	<b>3.28</b>	<b>1.68</b>	<b>1.87</b>	<b>7.84</b>	<b>2.61</b>
$R_o/R_e$	<i>1.07</i>	<i>0.96</i>	<i>1.07</i>	<i>1.23</i>	<i>1.06</i>	<i>1.32</i>	<i>1.45</i>
Coefficient on 'surveyed abroad'	0.020 (0.098)		-0.281 (0.174)	-0.552 (0.252)	-0.235 (0.199)		

Robust standard errors in parentheses.

As above, we regress the log of last home-country wage on a set of dummies describing age, education, sex, and classification of residence as rural or urban.<sup>25</sup> We calculate the residuals from these regressions and compare them for migrant versus nonmigrant heads of household and spouses.

Table 3 summarizes these calculations. The average residual for movers is in the 50<sup>th</sup> to 60<sup>th</sup> percentile of the distribution of residuals for non-movers in Mexico, Guatemala, Nicaragua, Costa Rica, Dominican Republic, and Haiti. For Peru it lies roughly in the 70<sup>th</sup> percentile. The exponent of the coefficient on the 'migrant' dummy in these regressions is an estimate of  $\mu_h$ . Using equation (3), this and the estimate of  $R_o$  allows the construction of an estimate of  $R_e$  for each country. For all but one country,  $R_o$  is indeed an overestimate of  $R_e$ . The ratio  $R_o/R_e$  is less than 1.23 for five of the seven countries; it is 1.32 for Haiti and 1.45 for Peru.

<sup>25</sup> The age dummies are identical to those used in regression (1). A dummy is included for each single year of education between 0 and 24. The Peru and Haiti data contain only urban observations; the Guatemala data contain only rural observations. The regressions include a dummy signifying that the respondent was contacted in the destination country rather than in the origin country. The regressions also include the year of the last domestic wage and the square of that year, to account for any tendency of wages recalled from further into the past to systematically differ from those recalled more recently. Age is measured at the time the wage was earned; education is measured at the time of the survey (but is unlikely to have changed greatly over time for the vast majority of these adults).

**Figure 4: Wage-history evidence on migrant selection: kernel densities of log home-country predicted wage for 35 year-old urban male with 9 years of schooling**

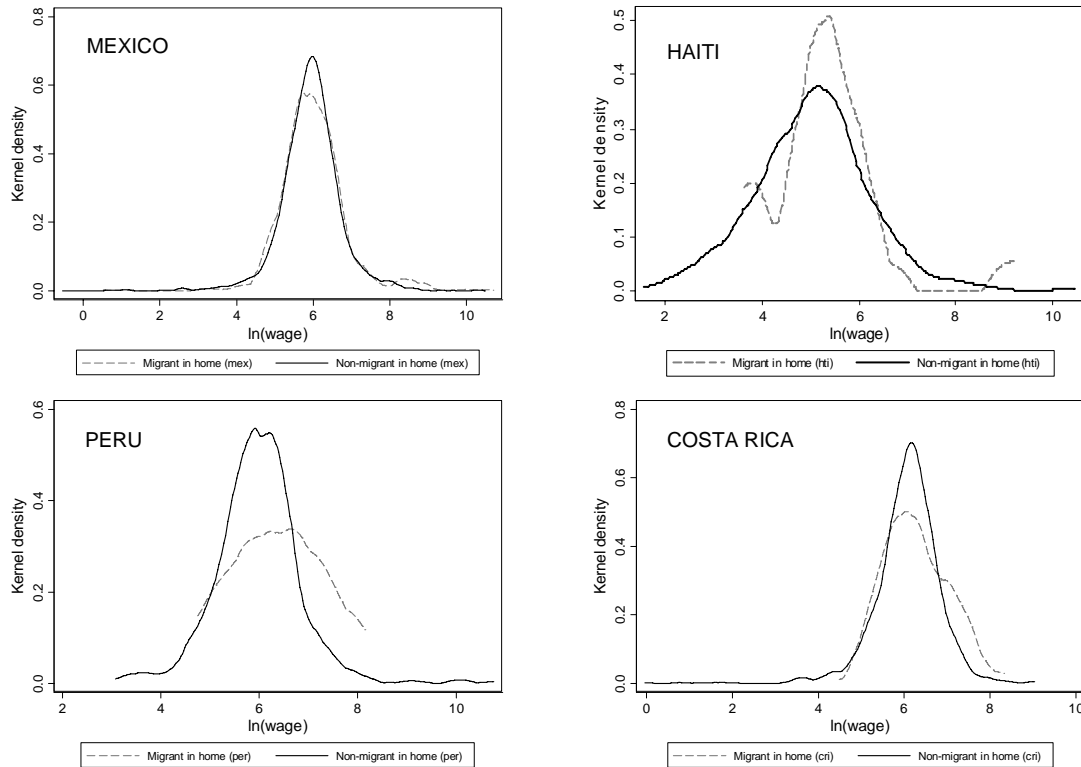


Figure 4 presents kernel density plots of the log wage regression residuals for four countries in Table 3. Each plot compares the distributions of unobservable home-country wage determinants among migrants and non-migrants.

This method has two drawbacks. First, many of the migrant interviewees are contacted in the country of origin, meaning that they are return migrants whose last observed wage in the home country occurs after migration, not before migration. If there is negative selection of return migrants this could exert downward bias on the estimates of the degree of positive selection. In four of the countries, however, substantial numbers of the interviewees are contacted in the destination country (row 3 of Table 4), meaning that they are not return migrants and their last home wage must have occurred prior to migration. Each regression includes a dummy variable identifying these individuals, and the coefficient on this dummy (bottom of Table 4) is never positive and significant. This

suggests that any negative selection of return migrants does not substantially shape these estimates.

A second drawback is that the survey samples are not nationally representative. The Mexican Migration Project and the Latin American Migration Project by an “ethnosurvey” method, in which target communities are chosen as subjectively representative of common migration processes by knowledgeable fieldworkers, and only within these communities are samples statistically representative. Massey and Zenteno (2000), however, compare a wide variety of household traits in data gathered by the Mexican Migration Project to data from a rigorously nationally representative survey of Mexico and find that biases introduced by the ethnosurvey method are “substantively unimportant.” Nevertheless, the estimates in Table 4 are better considered guidance regarding reasonable priors about the degree of migrant selection on unobservables, rather than definitive point estimates of the degree of selection.

### 3.5 *Existing microeconomic evidence on selection*

Only three existing studies of which we are aware estimate the precise degree of selection on unobserved determinants of earnings for emigrants from developing countries.

The only “experimental” evidence we know of is McKenzie, Gibson, and Stillman (2009) which compares wages of actual movers from Tonga who were randomly chosen from a pool of those who applied and were eligible for temporary employment in New Zealand. Their data and method allow estimate of  $R_o$  and  $R_e$ , and find positive selection on observables such that  $R_e$  is 4.91 while  $R_o$  is 6.14 and hence  $R_o/R_e = 1.25$ .<sup>26</sup>

---

<sup>26</sup> They take study New Zealand’s Pacific Access Category residence visa, which is designed to allow a limited number of citizens of Tonga (and three other island states) to settle in New Zealand each year via a random lottery. Any person age 18-45 who is a citizen and natural of the four PAC countries may register for the lottery, and among those registered a certain number are randomly allocated the chance to apply for residence. In their sample, the mean weekly income of Tongan non-applicants to the lottery is NZ\$70. The OLS estimate of the income gain to migration—which controls only for education, age, sex, height, and birth on Tonga’s principal island—is NZ\$360 per week. The experimental estimate is NZ\$274. Controlling for observables, then, the wage ratio  $R_o$  is  $(70+360)/70 = 6.14$ , while the true wage ratio  $R_e$  controlling for both observables and unobservables is  $(70+274)/70 = 4.91$ . In other words, the predicted wage ratios of observational equivalent workers overstates the true ratio of selected movers by a factor of  $6.14/4.91 = 1.25$ .

Gould and Moav (2008) (Figure 4) use nationally-representative panel data from Israel to suggest that emigrants are typically drawn from somewhat below the 50<sup>th</sup> percentile of the unobserved determinants of earnings—controlling for age, education, ethnicity, and native status. While the data they present do not allow an exact calculation of this percentile, they show that emigration rates are higher in the percentiles of the distribution of unobserved wage determinants below 50 than above 50.<sup>27</sup>

Fernández-Huertas Moraga (2008), using a panel of nationally-representative household survey data from Mexico, finds that movers come from just below the 50<sup>th</sup> percentile of the unobserved determinants of earnings. This suggests that the  $R_o$  for Mexico slightly underestimates  $R_e$ . For the present purpose this is very similar to the findings on Mexico we report.

These estimates agree well with our new evidence above and the three together suggest a ratio  $R_o/R_e$  between 0.9 and 1.25.

### 3.6 *Macroeconomic evidence*

Another way of assessing the bias induced by selection of migrants is to compare our individual based estimates with aggregate estimates. A core question in the economic growth literature is how much of the observed income differentials across countries are due to differences in the accumulation of factors—physical and human capital versus country specific productivity (e.g. Hall and Jones (1999), Caselli (2005)). There are two important implications of this literature for our results.

---

<sup>27</sup> Akee (2007b) finds some degree of positive selection on unobservables for emigrants from Micronesia, but the data presented in the paper do not allow calculation of where this places them in the distribution of the unobserved determinants of earnings for nonmigrants. There is additional, but less reliable evidence comparing destination-country incomes among migrant streams that plausibly differ in their basis and degree of selection. Cortes (2004) finds that refugees who arrived in the US between 1975 and 1980 had 6% lower earnings in 1980 than economic migrants who arrived in the same period—controlling for education, age, language ability, marital status, and region of residence. This means that either (i) if refugees are much less positively selected on unobservables than other migrants, then the selection bias in our estimates of  $R_e$  is not large, and is not far above 1.1, or (ii) refugees are selected on unobservables as much as other migrants, in which case their earnings relative to other migrants are not informative about the degree of bias that selection produces in our estimates.



First, if our wage ratios comparing observably identical workers are biased upward by the positive selection of migrants, then one would expect our estimates to be typically much higher than the macroeconomic estimates of the relative productivities of adjusted human capital, which are much less subject to this bias. Hall and Jones (1998) and Hall and Jones (1999) estimate a decomposition of countries' relative output per worker relative to the US into physical capital stocks, human capital, and country-specific productivity. The agreement is striking between our wage ratio  $R_o$  and the Hall and Jones growth accounting estimates of the relative marginal product of human-capital-equivalent workers for the 37 countries which have both. For these countries our median wage ratio  $R_o$  estimate in Table 1 is 3.82 and the mean (less four countries) is 4.53. The median estimate of the ratio of marginal products of human capital adjusted labor from Hall and Jones (1999) is 3.07, and the mean 3.92.

That is,  $R_o$  typically overestimates the Hall and Jones ratio by a factor of 1.25 (in medians) or 1.16 (when averages are taken without four outliers)<sup>28</sup> or at most 1.38 with means for all countries. This macro/micro gap, which is consistent with the macroeconomic differences not being affected by self-selected migrants, corresponds closely to the typical degree of selection bias we report in the preceding sections. This close general agreement across 37 countries is particularly striking as they are calculated by completely unrelated methods from completely unrelated data.<sup>29</sup>

---

<sup>28</sup> The four outliers, Egypt, Jordan, Venezuela, and Yemen, are countries for which Hall and Jones suggest the human capital adjusted gaps are much smaller than our estimated wage gaps, including estimates, in part because the HJ estimates are that total factor productivity in Yemen and Jordan is roughly equal to that in the US.

<sup>29</sup> The rank correlation between the two estimates is .41. The raw correlation without the four outliers mentioned above and Uganda is .75.

**Table 5: Comparison of estimated wage ratios of observably equivalent workers to Hall and Jones growth accounting estimates of the relative marginal product of human capital adjusted labor for 37 countries**

Row		Median	Mean (w/o four outliers)	Mean
I	$R_o$ estimates from Column 6, Table 1 for the 37 overlapping countries	3.82	4.53	5.11
II	Ratios of marginal product of human capital equivalent labor, US to country	3.07	3.92	3.69
III	Ratio of row I (preferred $R_o$ estimate) to row II (growth accounting estimate)	1.25	1.16	1.38
IV	Ratio of United States $A$ to country $A$ from Hall and Jones	2.44	2.90	2.71
V	Proportion of cross-national difference in human capital adjusted labor due to differences in $A$ (ratio IV to II)	0.80	0.74	0.74

Sources: Author's calculations and Hall and Jones (1998) (Table 7), Hall and Jones (1999).

Our micro-data based waged numbers also agree closely with the macroeconomic estimates of Hendricks (2002), who finds that even after controlling for cross-country differences in physical capital and both observable and unobservable human capital, GDP per capita in the United States is three times higher than in the average low-income country and eight times higher than in the poorest countries.<sup>30</sup>

A second important implication of these results is that the large majority of the cross-national gap in marginal products of workers with equivalent human capital is due to generalized productivity differences, not physical capital. This has consequences for the expected welfare gains of migration for people in the destination country: If the differences are mostly productivity (“ $A$ ”) and country-specific “ $A$ ” is purely nonrival, then there is little “factor shallowing” effect that would reduce wages for all existing workers by lowering the capital-labor ratio with labor inflows.

In fact, row III of Table 6 shows that the typical ratio of factor productivities ( $A$ ) is 2.44, which is roughly 80 percent of the observed wage (or marginal product) ratio. Caselli (2005) reviews the literature on growth decompositions and shows that, in the standard

<sup>30</sup> Hendricks (2002) compares earnings of observably identical workers from different countries in the United States to estimate the unobservable portion of human capital across countries, but his analysis is otherwise macroeconomic.

models, it is typical for physical and human capital differences to account for less than 50% of differences in per-worker output. As Easterly (2004) points out, in “productivity world” factors move to higher productivity locations, as opposed to “factor world” in which places with scarce factors attract more factors. In “productivity world” the gains to movers are not offset by losses to existing residents.

### 3.7 *Selection and migrant outcomes at the destination*

Immigrants’ wage outcomes at the destination may contain information about the degree of selection from the origin. If an immigrant typically earns a different wage than the observably identical native, this is the net effect of two conflicting forces: 1) the unobservable traits (e.g. “pluck” or ambition or entrepreneurial ability) that would have made the mover also earn higher wages than the observably identical person back home, and 2) how transferable those traits are from one country to another. Basically, if the migration process were producing people with tremendous “pluck” then, unless pluck evaporates at the border or there are large wage costs to being foreign-born in the US labor market, movers should tend make *more* than observably identical US residents.

Suppose that  $\ln \bar{w}_d = \ln w^* + (\theta \ln \mu_h - \ln \nu)$ , where  $\bar{w}_d$  is again the average wage of an immigrant at the destination controlling for observable traits, and  $w^*$  is the average wage of a native at the destination controlling for the same observable traits.  $\theta$  reflects the degree to which having above-average wages for one’s observable traits at the origin translates into having above-average wages at the destination. A low  $\theta$  means that most of what makes one person in the origin country earn more than the average of observably identical people “evaporates” when that person moves, and a high  $\theta$  means that it is mostly preserved when that person moves. The constant  $\nu$  captures any overall disadvantage of being a migrant: fewer connections, language difficulties, tacit knowledge about the job market, and so on. This gives

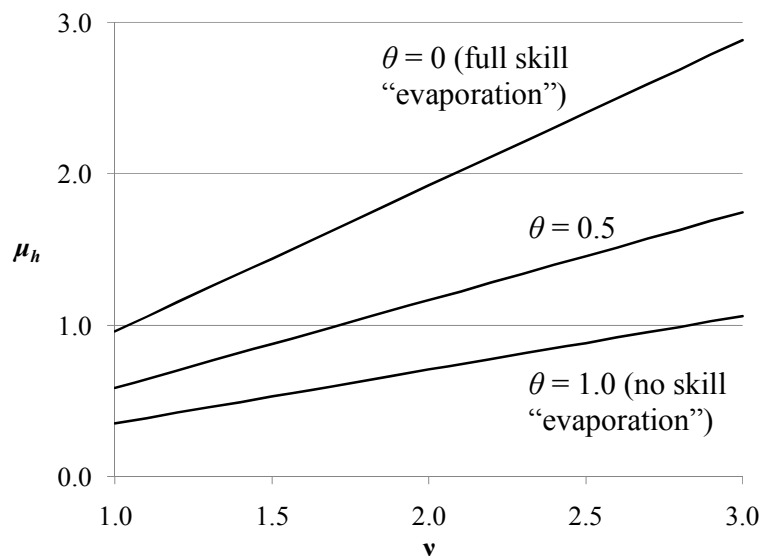
$$\mu_h = \frac{\nu}{e^\theta} \frac{\bar{w}_d}{w^*}. \quad (4)$$

Figure 2 suggests that immigrants' distributions of unobservable wage determinants do not greatly differ from those of natives; the ratio  $\bar{w}_d/w^*$  is not typically far from unity. Indeed, across the 42 countries, the mean ratio is 0.97, standard deviation 0.024, median 0.97, minimum 0.92 (Nepal), and maximum 1.03 (South Africa). As Hendricks (2002) observes in his "correction" of cross-national estimates of human capital for education quality differentials, the fact that migrant earnings in the US do not greatly differ from native-born earnings suggests that either 1) positive selection  $\mu_h$  among low to moderate skilled migrants is modest or 2) there is massive evaporation of unobserved wage determinants at the border.

We can take this a step further and estimate selection based on different assumptions about the evaporation of unobservable wage determinants and the costs of being foreign. Bolivians in the United States are close to the average as  $\bar{w}_d/w^* = 0.96$ . Figure 6 combines the various elements of equation 4 to yield implications about selection  $\mu_h$ . If the unobservable determinants of wages are partially transferable across borders ( $\theta = 0.5$ ), then the degree of selection at the origin does not rise above 1.5 unless the overall wage disadvantage to a Bolivian working in the US is such that a native earns close to triple the wage of an otherwise (observably *and* unobservably) identical Bolivian ( $\nu = 3$ ).

With a less extreme assumption about the inherent disadvantage of being Bolivian in the US labor market, say  $\nu = 1.5$ , then selection at the origin does not rise above 1.5 even if there is no international transferability of unobserved wage determinants at all ( $\theta = 0$ ). The fact that moderately skilled migrants make roughly the same as natives conditioned on observed traits (observed ratios  $\bar{w}_d/w^*$  between 0.95 and 0.99), appears to be, then, incompatible with sufficient positive selection to produce  $R_o/R_e$  far above 1.5.

**Figure 6: Selection and migrant outcomes at the destination**



### 3.8 Selection and migrant outcomes by occupation

Comparing international wage gaps within low-skill occupations offers still another way to assess the degree to which selection could cause  $R_o$  and  $R_e$  to differ. In the lowest-skill occupations, there is less scope for unobservable ability to substantially affect wages.

Previous researchers have documented that wage gaps across countries are enormous even for workers in the same sector, such as manufacturing, or in the same narrowly defined low-skill occupations, such as carpenters, laborers, or bus drivers (World Bank (1995)).

Table 7 gives a sampling of these estimates from other sources for the countries also in our sample. The median wage ratios within low-skill occupations are not substantially lower than the median estimates of  $R_o$  from Table 1. The ratio of real wages in the US to those in India for the same low-skill occupation is somewhere between 5 and 14. Our estimate for  $R_o$  is 6.25. If  $R_o$  were driven far above  $R_e$  by highly entrepreneurial Indians experiencing large earnings in the US, wage ratios within narrow low-skill occupations would be much smaller than  $R_o$ , whereas, if anything, the gaps in Table 7 are larger than those in column 6 of table 1.

**Table 6: Previous estimates of the ratio of wages in the US to those in other countries (PPP adjusted), without controlling for individual traits**

Source	Freeman & Oostendorp <sup>31</sup>		Rama & Artecona <sup>32</sup>	UBS <sup>33</sup>	$R_o$ (Table 1, column 6)
Occupation	Carpenter	Laborer	Industry	Laborer	
Year	1995	1995	1990-94	2006	2000
Median	6.36	7.67	4.26	4.65	
N	12	11	28	13	
<i>Selected countries</i>					
Bolivia	6.15	6.37	5.32		5.03
India	9.15	7.67	5.32	14.16	6.25
Mexico	6.57		2.78	7.49	2.53
Nigeria		10.60			14.85
Turkey			1.99	2.97	2.68

*N* gives the number of countries in the source that 1) have data for both the country in question and the US, and 2) are one of the 42 countries studied in this paper. Blank cells indicate no data for that country.

Wage ratios within low-skill occupations contain further information about the degree to which the estimated wage ratios could be influenced by selection of return migrants. If there is strong negative selection of return migrants, those with the poorest wage outcomes at the destination—people who hoped to make it as entrepreneurs, but were never able to leave construction or low-skill service jobs—might depart the destination country unobserved. This would tend to make the estimated wage ratios overstate the true gains to the marginal migrant. If this were exerting strong upward bias on  $R_o$ , wage gaps within low-skill occupations should be much lower. Comparing Table 6 with column 6

<sup>31</sup> Freeman and Oostendorp (2005) calculate average monthly wage rates for male workers, in US dollars at Purchasing Power Parity, in 1995. “Carpenter” refers to ILO occupation code 88 (“construction carpenter”), and “laborer” refers to ILO code 90.

<sup>32</sup> Rama and Artecona (2002) calculate “industry” wages as: “Labor cost per worker in manufacturing in current US dollars per year. Includes male and female workers. Calculated as the ratio between total compensation and the number of workers in the manufacturing sector as a whole. Compensation includes direct wages, salaries and other remuneration paid directly by the employer; plus all employers’ contributions to social security programs on behalf of their employees. Data on labor costs per worker are from plant-level surveys covering relatively large firms, mostly in the formal sector of the economy. Figures are converted into US dollars using the average exchange rate for each year. In countries of the former Soviet Union, the exchange rate of 1989 is used for previous years.” “Government” wage is “Average wage of employees in the central or general government, in current US dollars per year. Includes male and female employees. Calculated dividing the government payroll by the total number of employees. Data are from government records. Figures are converted into US dollars using the average exchange rate for each year.” Both of these are converted to PPP dollars using the PPP-to-official-exchange-rate ratio from World Bank (2007).

<sup>33</sup> The UBS estimates (Hoefert and Hofer (2007)) are for urban areas (respectively: Buenos Aires, New Delhi, Seoul, Mexico City, Manila, Bangkok, and Istanbul, with the US represented by New York City), and show the hourly wage (assuming 50 working weeks per year) of a “building laborer”, 25 years old, single, unskilled or semi-skilled (p. 41) adjusted for cost of living in each city by the prices of 95 goods and 27 services (p. 8).

from Table 1 shows that this is not at all the case. The estimates of  $R_o$  are not systematically higher than wage gaps within low-skill occupations.

### 3.9 *Summary of evidence on selection*

Table 7 summarizes the evidence on the degree to which  $R_o$  might exceed  $R_e$ . That is, it summarizes the degree to which positive selection of moderate skill migrants on unobservable wage determinants causes wage ratios of observably equivalent workers to overstate the wage gap of equal productivity workers (and hence wage gain of a marginal mover). The new and existing evidence we discuss above suggests that the degree of positive selection of existing movers on unobserved wage determinants is modest at best, typically inducing a bias of 1.0 to 1.3, with the greatest observed bias (for Peru) at 1.45.

Table 8 shows what would be the consequences of scaling back the estimates of  $R_o$  from column 6 of table 1 to account for the range of selection typically observed, in two different ways. Columns 2, 4, 6 and 8 answer the question, “What would  $R_e$  be if migrants were drawn from the  $x^{\text{th}}$  percentile of the distribution of the wages of observably equivalent workers in the home country?” Columns 3, 5, 7, and 9 show the extent to which this differs from the column 6 of table 1, expressed as the ratio of  $R_o/R_e$ .

**Table 7: Summary of estimates of the bias of wage ratios of “observably identical” workers to “equal productivity” workers**

Countries	Data and method (source)	Estimated $R_o/R_e$	Estimated migrant percentile
Philippines	9-month panel, nationally representative, compare wage residuals of movers to non-movers when part of household remains behind.	1.10	54
Mexico	Same as above, but 12-month panel.	1.03	57
South Africa	4-year panel of young adults in Cape Town area, compare wage residuals of movers to non-movers.	1.20	55
Haiti, Costa Rica, Guatemala, Peru, Nicaragua, Mexico, Dominican Rep.	Retrospective wage histories, ethnosurvey method. Compare wage residuals of non-movers, former movers, and current movers.	0.96-1.45 (median 1.07)	47-69 (median 54)
Tonga (to NZ)	Experimental estimate of wage gain exploiting random visa lottery versus unadjusted estimates of wage gain (McKenzie, Gibson, and Stillman (2009)).	1.25	—
Mexico	Panel comparing the wage regression residual distribution of subsequent movers to non-movers (Fernández-Huertas Moraga (2008)).	<1	<50
Israel	Panel comparing the wage regression residual distribution of subsequent movers to non-movers (Gould and Moav (2007) (Figure 4)).	<1	<50
37 countries	Comparing our estimates of $R_o$ to Hall and Jones (1999) estimates from macroeconomic data of the ratio of marginal products of human capital adjusted labor.	1.16-1.38	
At median of 42 countries	Scenarios of selection based on observed $\bar{w}_d/w^*$ and supposed $\theta$ and $\nu$	$\nu = 2.0, \theta = 0.5$	1.17
		$\nu = 2.0, \theta = 0.2$	1.57
		$\nu = 1.5, \theta = 0.2$	1.18
Bolivia, India, Mexico, Nigeria, Turkey	Comparing $R_o$ to three different sources of wage gaps within low-skill occupations.	Roughly 1	



Column 8 shows what  $R_e$  would be if our comparison group of late arrivers in the US were drawn from the 70<sup>th</sup> percentile of origin country residual—which is stronger than *any* of the evidence from *any* country above supports. Even in this extreme case the median estimate of  $R_e$  is 3.4—equal productivity workers make more than triple in the US—and in 38 of 42 countries even under these assumptions about selection,  $R_e > 2$ .

The final column of table 8 illuminates the selection issue from a different angle, which is to calculate how strong positive selection on unobservable wage determinants would have to be in order for the wage ratio to be equal to 2. (In other words, workers' wages would double at the border and such a ratio, as we show below, is almost certainly higher than could be sustained as an equilibrium in the absence of policy barriers.) Looking at the distributions in Figure 2, this exercise ascends the residual distribution of home-country workers until the ratio with the mean of the “late arrivers” is just equal to 2. For nearly all countries the extent of positive selection even to produce this ratio is far higher than any of the diverse evidence from Table 7 supports—the median is the 90<sup>th</sup> percentile.

Figure 3 provides a graphical representation of the ratios  $R_e$  in Table 8, under different assumptions about the percentile of origin-country unobserved wage determinants from which migrants come.<sup>34</sup> Take, for example, the median Peruvian-born, Peruvian educated 35 year-old urban male with 9 years of education who resides in the US. If his counterfactual wage lies at the median of the distribution of wages for observably identical workers in Peru,  $R_e$  is 4.1. If his counterfactual wage lies at the 40<sup>th</sup> percentile,  $R_e$  is 4.8. If it is at the 60<sup>th</sup> percentile,  $R_e$  is 3.5, and if it is at the 70<sup>th</sup> percentile  $R_e$  is about 3. (This differs slightly from the estimate in Table 3, which reflects the mean migrant rather than the median migrant.)

---

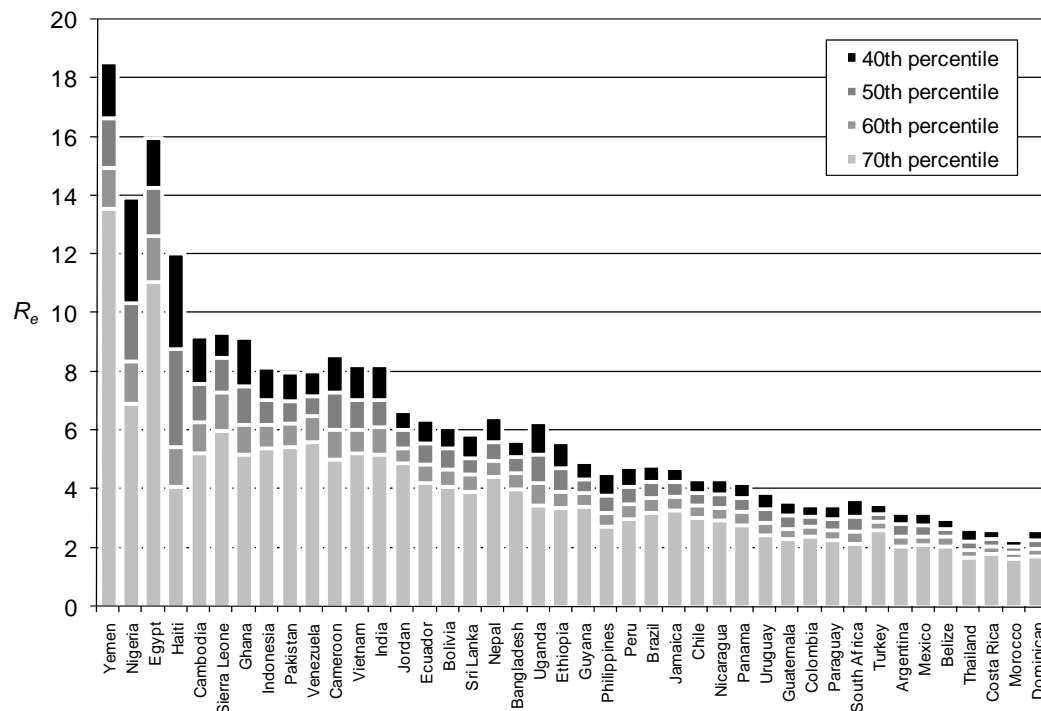
<sup>34</sup> These are nonparametric estimates in the sense that they impose no assumptions on the form of the distribution of the unobserved determinants of earnings in the migrants' home country nor the distribution in the destination country.

**Table 8: Implied  $R_e$  under different assumptions about selection**

Column:	Col 6, Table 1	Assumed percentile of median migrant wage in origin-country distribution of wages for observably identical workers								Origin percentile required if $R_e = 2$
	$R_o$	40 <sup>th</sup> percentile		50 <sup>th</sup> percentile		60 <sup>th</sup> percentile		70 <sup>th</sup> percentile		
	<b>(1)</b>	$R_e$	$R_o/R_e$	$R_e$	$R_o/R_e$	$R_e$	$R_o/R_e$	$R_e$	$R_o/R_e$	<b>(10)</b>
Yemen	15.45	<b>18.54</b>	0.83	<b>16.64</b>	0.93	<b>14.94</b>	1.03	<b>13.54</b>	1.14	98.5
Nigeria	14.85	<b>13.93</b>	1.07	<b>10.34</b>	1.44	<b>8.36</b>	1.78	<b>6.92</b>	2.15	99.2
Egypt	11.92	<b>15.95</b>	0.75	<b>14.28</b>	0.83	<b>12.62</b>	0.94	<b>11.06</b>	1.08	99.9
Haiti	10.31	<b>12.03</b>	0.86	<b>8.76</b>	1.18	<b>5.43</b>	1.90	<b>4.08</b>	2.53	85.9
Cambodia	7.45	<b>9.19</b>	0.81	<b>7.57</b>	0.98	<b>6.28</b>	1.19	<b>5.21</b>	1.43	96.5
Sierra Leone	7.43	<b>9.33</b>	0.80	<b>8.47</b>	0.88	<b>7.28</b>	1.02	<b>5.99</b>	1.24	99.6
Ghana	7.12	<b>9.14</b>	0.78	<b>7.50</b>	0.95	<b>6.20</b>	1.15	<b>5.17</b>	1.38	94.8
Indonesia	6.72	<b>8.14</b>	0.83	<b>7.05</b>	0.95	<b>6.20</b>	1.08	<b>5.39</b>	1.25	98.8
Pakistan	6.57	<b>7.96</b>	0.83	<b>7.00</b>	0.94	<b>6.23</b>	1.05	<b>5.43</b>	1.21	98.6
Venezuela	6.57	<b>8.01</b>	0.82	<b>7.16</b>	0.92	<b>6.47</b>	1.02	<b>5.61</b>	1.17	98.9
Cameroon	6.53	<b>8.58</b>	0.76	<b>7.30</b>	0.89	<b>6.03</b>	1.08	<b>5.02</b>	1.30	96.5
Vietnam	6.49	<b>8.23</b>	0.79	<b>7.03</b>	0.92	<b>6.04</b>	1.07	<b>5.22</b>	1.24	99.1
India	6.25	<b>8.24</b>	0.76	<b>7.05</b>	0.89	<b>6.09</b>	1.03	<b>5.16</b>	1.21	97.0
Jordan	5.65	<b>6.68</b>	0.85	<b>6.02</b>	0.94	<b>5.40</b>	1.05	<b>4.88</b>	1.16	99.3
Ecuador	5.16	<b>6.37</b>	0.81	<b>5.54</b>	0.93	<b>4.86</b>	1.06	<b>4.22</b>	1.22	97.4
Bolivia	5.03	<b>6.10</b>	0.82	<b>5.37</b>	0.94	<b>4.66</b>	1.08	<b>4.09</b>	1.23	95.0
Sri Lanka	4.95	<b>5.87</b>	0.84	<b>5.07</b>	0.98	<b>4.49</b>	1.10	<b>3.90</b>	1.27	97.0
Nepal	4.85	<b>6.46</b>	0.75	<b>5.60</b>	0.87	<b>4.95</b>	0.98	<b>4.42</b>	1.10	98.2
Bangladesh	4.60	<b>5.65</b>	0.81	<b>5.11</b>	0.90	<b>4.53</b>	1.02	<b>4.01</b>	1.15	97.1
Uganda	4.38	<b>6.28</b>	0.70	<b>5.20</b>	0.84	<b>4.22</b>	1.04	<b>3.45</b>	1.27	91.0
Ethiopia	4.35	<b>5.61</b>	0.78	<b>4.70</b>	0.93	<b>3.93</b>	1.11	<b>3.35</b>	1.30	90.4
Guyana	3.87	<b>4.94</b>	0.78	<b>4.34</b>	0.89	<b>3.85</b>	1.01	<b>3.40</b>	1.14	93.5
Philippines	3.82	<b>4.56</b>	0.84	<b>3.77</b>	1.01	<b>3.21</b>	1.19	<b>2.73</b>	1.40	87.3
Peru	3.79	<b>4.77</b>	0.79	<b>4.08</b>	0.93	<b>3.47</b>	1.09	<b>2.98</b>	1.27	90.1
Brazil	3.76	<b>4.80</b>	0.78	<b>4.23</b>	0.89	<b>3.69</b>	1.02	<b>3.20</b>	1.18	91.0
Jamaica	3.63	<b>4.72</b>	0.77	<b>4.23</b>	0.86	<b>3.73</b>	0.97	<b>3.29</b>	1.10	92.5
Chile	3.53	<b>4.34</b>	0.81	<b>3.89</b>	0.91	<b>3.45</b>	1.02	<b>3.04</b>	1.16	91.9
Nicaragua	3.52	<b>4.32</b>	0.81	<b>3.83</b>	0.92	<b>3.36</b>	1.05	<b>2.92</b>	1.21	87.7
Panama	3.36	<b>4.22</b>	0.80	<b>3.69</b>	0.91	<b>3.25</b>	1.03	<b>2.79</b>	1.20	86.1
Uruguay	3.10	<b>3.86</b>	0.80	<b>3.32</b>	0.93	<b>2.86</b>	1.08	<b>2.44</b>	1.27	80.5
Guatemala	2.94	<b>3.58</b>	0.82	<b>3.11</b>	0.95	<b>2.65</b>	1.11	<b>2.30</b>	1.28	77.2
Colombia	2.88	<b>3.46</b>	0.83	<b>3.08</b>	0.94	<b>2.71</b>	1.06	<b>2.37</b>	1.22	83.3
Paraguay	2.78	<b>3.44</b>	0.81	<b>3.00</b>	0.93	<b>2.62</b>	1.06	<b>2.25</b>	1.24	78.3
South Africa	2.75	<b>3.67</b>	0.75	<b>3.08</b>	0.89	<b>2.56</b>	1.07	<b>2.12</b>	1.30	73.7
Turkey	2.68	<b>3.49</b>	0.77	<b>3.17</b>	0.85	<b>2.88</b>	0.93	<b>2.58</b>	1.04	86.6
Argentina	2.54	<b>3.20</b>	0.79	<b>2.81</b>	0.90	<b>2.41</b>	1.05	<b>2.07</b>	1.23	77.1
Mexico	2.53	<b>3.19</b>	0.79	<b>2.79</b>	0.91	<b>2.40</b>	1.05	<b>2.08</b>	1.22	72.2
Belize	2.43	<b>3.00</b>	0.81	<b>2.64</b>	0.92	<b>2.38</b>	1.02	<b>2.03</b>	1.20	71.5
Thailand	2.17	<b>2.64</b>	0.82	<b>2.24</b>	0.97	<b>1.94</b>	1.12	<b>1.65</b>	1.32	66.1
Costa Rica	2.07	<b>2.58</b>	0.80	<b>2.29</b>	0.90	<b>2.03</b>	1.02	<b>1.79</b>	1.16	61.1
Morocco	2.00	<b>2.28</b>	0.88	<b>2.04</b>	0.98	<b>1.83</b>	1.09	<b>1.62</b>	1.23	51.1
Dom. Rep.	1.99	<b>2.58</b>	0.77	<b>2.26</b>	0.88	<b>1.98</b>	1.01	<b>1.71</b>	1.16	59.2
<b>Median</b>	<b>4.11</b>	<b>5.28</b>	<b>0.805</b>	<b>4.52</b>	<b>0.92</b>	<b>3.89</b>	<b>1.05</b>	<b>3.38</b>	<b>1.225</b>	<b>91.45</b>

Table reflects predicted wages for 35 year-old urban male with 9 years of education. The first column reproduces column 6 of Table 1. "Residual wage" refers to the residual when ln wage is regressed on education, age, sex, and rural/urban as described in the text. The rightmost column shows the percentile of the origin-country distribution of unobserved determinants of income from which the mean migrant would need to be drawn in order for  $R_e$  to equal 2 given the estimated  $R_o$ .

**Figure 3: Estimated  $R_e$  under different assumptions about selection**



Countries are sorted left to right by the estimate of  $R_e$  in Table 1, column 6. Different shades show assumed expected percentile of the marginal migrant in the distribution of the unobservable determinants of earnings in the migrant’s home country.

#### 4 Interpreting the estimates: To what degree are wage gaps sustained by policy barriers?

The preceding section gives several reasons to believe that the wage ratios in Table 1 are modest overestimates of the “place premium”—wage ratios for workers of equal intrinsic productivity—but that there are nevertheless enormous gains in real consumption wages from moving across the border. It is at least conceivable that the estimated gains in wages are sufficiently offset by transport costs, credit constraints, and by other psychic costs of movement across borders so that the wage gaps are not policy induced but reflect compensating differentials in equilibrium.

There is, however, strong evidence that border restrictions do prevent movement. The Gallup organization surveyed individuals around the world and asked: “Ideally, if you

had the opportunity, would you like to move permanently to another country, or would you prefer to continue living in this country?” For the 38 countries for which our data overlap, fully 25 percent of individuals reported they would like to live in another country, including 42 percent of all Peruvians, 42 percent of all Bangladeshis, 46 percent of all Ethiopians, and 58 percent of all Nigerians, but only 9 percent of Indonesians or Indians.<sup>35</sup>

Even these numbers likely understate the demand for generalized labor mobility, as they asked only about *permanent* movement. A survey conducted for the World Bank’s World Development on Youth asked individuals aged 15-24, “If it were possible for you to legally move to another country to work, would you?” with options of “move permanently”, “move temporarily”, “try it out”, and “not move”. Two of the surveyed countries overlap with our estimates and the Gallup Poll: Bangladesh and Ethiopia. In Bangladesh the expressed demand for temporary movement among males was enormous; only about 10 percent said they would “not move” but only 5 percent would “move permanently”—about 75 percent said they would move temporarily or try it out. In Ethiopia, only about a quarter said they would “not move” and again most of the demand for movement was temporary or “try it out” (World Bank (2007) (Figure 8.5)). While these surveys are difficult to compare because youth are more likely to move than older individuals and the sample frames are not the same, it does suggest that demand for temporary labor mobility is large compared the desire for permanent migration. In either case, while this evidence is not about moving to the US and we do not assert that the only motivation of movement is wage gains, nevertheless the expressed demand for movement “ideally” or “if it were legal” is many times, if not orders of magnitude, larger than actual cross-border movement.

We now turn to estimating what fraction of the wedge  $R_e$  might be reasonably believed to be sustained by policy barriers to labor movement. We approach this question by estimating wage ratios in settings where there are few legal barriers to movement and migration is entirely self-selected: (i) movement from Puerto Rico, Guam, and

---

<sup>35</sup> <http://www.gallup.com/poll/108325/OneQuarter-Worlds-Population-May-Wish-Migrate.aspx>.

Micronesia to the United States; (ii) movement between US states; (iii) movement from French Overseas Departments to metropolitan France; and (iv) movement between Europe and the New World before the establishment of substantial policy barriers at the destination.

#### 4.1 *The US zone of free movement*

Certain states and territories associated with the US provide a useful test case as they are geographically and culturally similar to countries in the sample, but face almost no policy barrier against movement to the continental US. People from the Associated Free State of Puerto Rico, for example, enjoy US citizenship allowing free movement to the continental US. People from the Territory of Guam likewise hold US citizenship and can move at will to the continental US.

Table 9 begins by using data from the US census to carry out regressions identical to equation (1) to estimate  $R_o$  for Puerto Rico and Guam.<sup>36</sup> Here, “foreign” born signifies birth in Puerto Rico or Guam, rather than the mainland. Analogously to Table 1, the first two columns do not control for any individual traits. Column 1 places mainland-born mainland-residents in the numerator, and column 2 places “foreign”-born mainland residents in the numerator. Columns 3 and 4 repeat the previous columns but include the same controls for age, education, sex, and rural/urban residence from regression (1). As before, the wages are adjusted for purchasing power.<sup>37</sup> The first row of ratios compares people in the mainland US to people in Puerto Rico, and the second row restricts the comparison to people who describe themselves on the census form as ethnically Puerto

---

<sup>36</sup> With the exception that all foreign-born US residents are combined in a single “foreign”-born, US-resident category, as the census does not report year of arrival on the mainland.

<sup>37</sup> Dollar wages in Puerto Rico are converted to “mainland PPP” dollars by dividing by the PPP factor 0.86 from Heston, Summers, and Aten (2006). Dollar wages in Guam are converted to mainland purchasing power using figures from the Office of Personnel Management (2000: 44142) of the US, which determined in a 1998 survey that a basket of basic consumption items (including food at home and away, tobacco, alcohol, furnishings, clothing, domestic services, professional services, personal care, and recreation) cost about 20.3% more in Guam than in Washington, DC. This figure is not sensitive to different spending patterns at different income levels.

Rican. The third row compares people in mainland US to people in Guam, and the fourth row restricts the comparison to ethnically Guamanian (Chamorro) people.<sup>38</sup>

**Table 9: Wage gaps across observably identical workers without policy barriers**

Numerator	(1)	(2)	(3)	(4)	No. of “foreign”-born observations	
					On main-land	Off main-land
US-born	X	—	X	—		
“Foreign” born	—	X	—	X		
Controls?	—	—	X	X		
Puerto Rico	1.63 (1.60, 1.65)	1.42 (1.39, 1.46)	2.02 (1.94, 2.10)	1.55 (1.46, 1.65)	30,900	47,085
Puerto Rico (ethnic Puerto Ricans only)	1.36 (1.34, 1.38)	1.42 (1.40, 1.44)	1.72 (1.66, 1.78)	1.54 (1.49, 1.60)	27,906	46,640
Guam (all races)	1.50 (1.36, 1.65)	1.40 (1.24, 1.58)	1.55 (1.22, 1.97)	1.55 (1.14, 2.11)	1,923	2,406
Guam (Chamorro only)	1.34 (1.25, 1.43)	1.37 (1.27, 1.48)	1.14 (0.89, 1.45)	1.34 (1.11, 1.62)	667	1,262
Micronesia				~1.1–1.4		

95% confidence interval for ratio estimates shown in parentheses

The results suggest that the average Puerto Rican born, ethnically Puerto Rican 35 year-old urban male with 9 years of education earns about 1.5 times as much as the observably identical person in Puerto Rico. For ethnic Guamanians born on Guam, the ratio is 1.3.

Table 8 lists another, less-studied case of a developing country facing no migration barrier from the US. Since 1986, any citizen of the Federated States of Micronesia (FSM) can acquire on demand a US work visa of unlimited duration. Akee (2007b) presents linked microdata on 632 individuals who were in FSM at the time of the 1994 census and had migrated to the United States by the time of the Micronesian Immigrant Survey of Hawaii and Guam in 1997. Many people in the sample were self-employed or unemployed prior to migration, and only 92 report wage income before and after

<sup>38</sup> “Chamorro” people are identified in Guam as people who indicate “ethnicity” as “Pacific islander” *and* indicate that they speak the Chamorro language. “Chamorro” are identified in the US as having been born in Guam and listing “race” as “Guamanian or Chamorro alone”.

migration. Mean annual pre-migration earnings are roughly US\$4,000 and mean annual post-migration earnings are roughly US\$8,000.<sup>39</sup> Using the PPP deflator of 0.51 for FSM with respect to the US as a whole (Heston, Summers, and Aten (2006)), purchasing power-adjusted wages would be roughly the same before and after migration. But group housing arrangements, food choices, and other strategies allow recently-arrived FSM migrants to live at somewhat less than the typical cost of living for full US citizens, suggesting that the ratio representing their true real wage gain to migration is likely to be in the range of 1.1 to 1.4. Note that in this case the same individuals are being followed across the border, so selection bias is absent. This suggests once again that even in the presence of substantial transportation costs, wage gaps induced by natural barriers alone are limited across spatially and culturally distinct but legally integrated labor markets.

We can compare these gaps to those observed within a spatially, culturally, *and* legally integrated market: The continental US. In the US census data, we estimate wage gaps for observably identical workers across the 48 continental United States (without controlling for state of birth or state of education), with New York in the denominator. Log wages adjusted for purchasing power<sup>40</sup> are regressed on the same dummies for age, years of education, sex, and rural/urban residence as in regression (1)<sup>41</sup> and these regressions predict wages for a 35 year-old urban male with 9 years of education, as before. The highest predicted wage relative to New York is found in Texas (1.18), the lowest in North Dakota (0.85). The highest ratio sustained across any of the 48 states is, then, the ratio of highest to lowest: 1.38.

These results for US states accord well with domestic rural-urban wage gaps for observably identical workers within all 43 developing countries we study. The average coefficient on the “rural” dummy for foreign residents in the log wage regression (1) is approximately  $-0.3$ . This suggests that the average ratio of an urban worker’s wages to

---

<sup>39</sup> Wage and price information in personal communication from the author. These data only capture a relatively short window of experience in the destination (less than two years) and wages may increase over a longer time period.

<sup>40</sup> Dollar wages in each of the US states are standardized for purchasing power based on the state cost-of-living estimates by Berry, Fording, and Hanson (2000). Only 48 states are estimated because the cost-of-living estimates used to standardize wages by purchasing power are missing for Alaska and Hawaii (and the District of Columbia). The state of New York is used as the base group.

<sup>41</sup> “Urban” is defined as residing within any of the census-defined “metropolitan” areas of the country.

those of an otherwise observably identical rural worker is about 1.4 across all these countries on average.<sup>42</sup> Such a wedge reflects the combined effect of natural barriers and selection bias. In legally integrated labor markets, movement itself places bounds on how high rural/urban wage ratios can rise without population movements.

#### 4.2 *The French zone of free movement*

Guadeloupe, Martinique, French Guyana, and Réunion are overseas departments of France. They are also essentially developing countries with very different geographic, cultural, ethnic, economic, and historical traits from metropolitan (continental European) France. People born there are French citizens and face no policy barrier if they wish to move to the *métropole*. The French government calculates that in 1999, average household disposable income per household member in the *métropole* was 1.59 times its average across the four overseas departments.<sup>43</sup> This can be considered a PPP ratio because basic consumption prices in the overseas departments are very similar to those in the *métropole*.<sup>44</sup> Unfortunately, no publicly-available microdata allow us to repeat the exercise in Table 8 for France by controlling for individual traits. But populations in the overseas departments are generally younger and less educated than in the *métropole*, so this ratio would certainly be lower if individual traits were controlled for. We can thus consider 1.6 to be an approximate upper bound on  $R_o$  in the French zone of free movement.

#### 4.3 *The Atlantic economy before 1914*

Prior to the introduction of a variety of legal restrictions following the first World War, emigrants from Europe to the New World faced few policy barriers to movement.

---

<sup>42</sup> This is in line with the estimates in Sir Arthur Lewis's (1954) Nobel-winning research, in which he observed that rural-urban migration in developing countries is generally triggered when urban-rural real wage gaps rise substantially above 1.3.

<sup>43</sup> Ratio of household disposable income per household member in provincial metropolitan France (€16,697) to the same figure in the Overseas Departments (€10,482). "Revenu disponible brut des ménages par région (RDB)", from INSEE (2000), *Comptes économiques régionaux des ménages - base 2000*, <<http://www.insee.fr/fr/ffc/figure/CMRSOS04238.xls>>, accessed September 22, 2008.

<sup>44</sup> INSEE Antilles-Guyane (2008), *L'Indice des Prix*, No. 11, December (Point-à-Pitre: INSEE Service Régional de Guadeloupe), <[http://www.insee.fr/fr/themes/document.asp?reg\\_id=26&ref\\_id=13474](http://www.insee.fr/fr/themes/document.asp?reg_id=26&ref_id=13474)>, accessed September 22, 2008.



Williamson (1995), Williamson (1997) and Hatton and Williamson (2005) show that large-scale migration during this period contributed to tremendous convergence of low-skill wages. Using Williamson's data,<sup>45</sup> Figure 7 shows the ratio of real urban construction wages in the United States to the wage in several major migrant-origin countries in the years up to 1913. By the end of that period the ratio for all of these countries except Italy had fallen below 1.8. Italy's migration started much later and was plummeting in 1913, so this can be considered an out-of-equilibrium ratio.

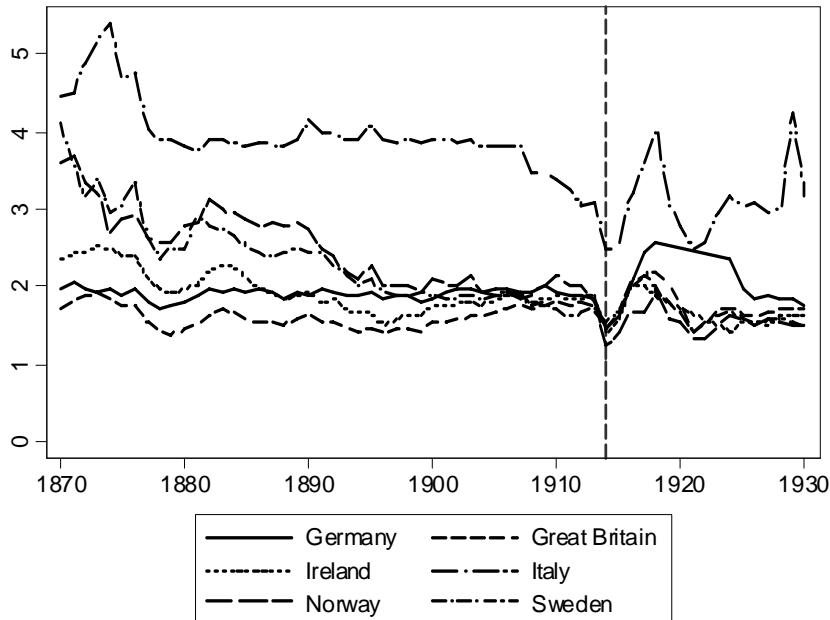
Williamson (1999) (Figure 9) shows that a very similar pattern developed between migrant-origin countries of the southern Mediterranean and destination countries in Latin America. He constructs, for the decades leading up to 1913, the ratio of real urban unskilled wages in Argentina, Colombia, Cuba, Mexico, and Uruguay to a combined wage index for Italy, Portugal, and Spain. This ratio fell from between 2 and 3 around 1880 to an average of about 1.7 around 1913, as large-scale emigration to the New World proceeded. Together with the North American evidence above, this suggests that even at a time when transportation, information, and (arguably) credit costs were substantially higher than today's, and even in the presence of substantial language and cultural barriers, wage ratios above 1.8 could not be sustained in the absence of policy barriers to movement.

The lesson of this evidence is that in areas that lack policy barriers to labor movement, we do not observe wage ratios for observably equivalent workers across space that exceed 1.8 in any historical period, or about 1.5 currently. We have shown that wage ratios even for fully equivalent workers from Haiti, the Philippines, Nicaragua, and Peru lie far above 1.5-1.8. Table 8 strongly suggests that the same is true for Egypt, Nigeria, Pakistan, Vietnam, and other countries. We interpret this to mean that the very large wage ratios we observe for many countries are sustained by policy barriers to movement.

---

<sup>45</sup> Kindly provided by the author.

**Figure 7: Ratio of US real urban construction wage to that of major migrant origin countries, 1870-1913**



## 5 Comparisons of wage gaps and policy implications

Economists often compare the existence of market inefficiencies, were they to exist, to “dollar bills lying on the sidewalk”: Price differences create profit-making opportunities. Barriers to labor mobility have created differences in the price of moderate-skill labor such that there are ten thousand dollar bills lying on the sidewalks in the US. We compare the distortions in labor markets to international price gaps for goods and factors. We also compare policy-induced wage discrimination at the border to other forms of wage discrimination. Finally, we compare the potential gains from migration to other forms of poverty reduction policies or programs.

### 5.1 Comparison with other border-induced price wedges

Trade economists have frequently used directly estimated price differentials of equivalent products to estimate the impacts of non-tariff barriers on goods and services (Anderson

and Neary (1994) *inter alia*). The wage ratios reported here use the price of equivalent labor, corrected for worker productivity (both observed and unobserved) and compensating differentials, to measure the price effects of the existing non-tariff restrictions on labor flows. Strikingly, there is a huge flourishing literature on the extent of “globalization” that debates just how “flat” the world might be for goods and capital that often takes the cliff created by border restrictions for workers for granted.

Many gaps in goods prices have been small for a century: O'Rourke and Williamson (2000) (pp. 29-55) report that the transportation revolution in the late nineteenth century brought prices of the most basic commodities, such as wheat, to near parity between the US and Great Britain. The remaining explicit barriers to trade in many manufactured goods are small in the OECD.

There has been some revisionism on the extent to which markets are truly “borderless”—detailing that even in quite liberal trade environments there is much less trade than would be expected, as McCallum (1995) documents for US-Canada trade. Detailed price comparisons for 120 industries in 1999 executed by Bradford and Lawrence (2004) (p. 7) reveal that consumer prices in many European countries are 30-50% higher than the lowest available prices for essentially the same goods in any country, while prices in Japan are roughly 100% higher and these differences persist even when one nets out retailing costs. Bradford and Lawrence point to these deviations from goods price equalization as evidence that globalization has not “gone far enough” as price differentials remain substantial. But the substantial gaps they find, of 50 to 100 percent (only in some commodities, not on average) are the smallest of the wage gaps.

The situation could not be more different in the market for capital, both in terms of prices of financial instruments as well as in marginal products. Lamont and Thaler's (2003) survey of the literature on cross-border price gaps for equivalent financial instruments suggests that price gaps of 15% (not percentage points) are considered “inexplicably large.” Casselli and Feyrer (2006) calculate the marginal product of capital across countries and find that marginal products of capital across countries are “essentially

equalized.” In fact, by their estimates, which correct the “naïve” estimates of the marginal product of capital for differences “natural capital” and the price of output, the return to capital is *lower* in poor countries than rich countries (8.4% vs. 6.9%, Table 2).

The fact that price gaps for labor exceed price gaps for goods, often by an order of magnitude, has obvious implications for the social welfare effects of liberalization as the heuristic intuition is that in simple partial equilibrium welfare losses increase with the square of price wedges. If one considers the general equilibrium welfare/output gains from completely open borders based the gains are astronomical: Hamilton and Whalley (1984) estimate a rough doubling of world output per person and Klein and Ventura (2004) use a calibrated general equilibrium model with capital mobility and estimate gains between 94% and 172%. But one need not consider radical notions like open borders, the general equilibrium simulations of Walmsley and Winters (2005) (Table 4, col. V) suggest that just a tiny relaxation of barriers to migration into the OECD—allowing an additional movement of 3% of the existing labor force—would raise the welfare of those moving by \$170 billion.<sup>46</sup> The World Bank (2005) (p. 128) finds that following an elimination of *all* remaining policy barriers to trade worldwide, developing countries would gain \$109 billion in annual income by 2015. Casselli and Feyrer (2006) estimate the welfare gains from *complete* equalization of the MPK at one tenth of one percent of world GDP—roughly \$65 billion—again a small fraction of the gains from only a modest relaxation of barriers to labor movement.

## 5.2 *Comparison with other forms of wage discrimination*

By “wage discrimination” we mean any labor market force that causes workers of equal intrinsic productivity involuntarily to receive different wages because of socially constructed characteristics of the workers. Barriers to the movement of workers across international borders create wage discrimination of precisely this kind. Suppose a given

---

<sup>46</sup> Their simulations were based on rough assumptions about how much of the existing wage differences were due to productivity differences that would move with the worker (which they assume is only half) while our estimates per worker are based on data. They find a welfare gain to movers of roughly \$20,000 per mover, which is close to our estimate for India of P\$19,900 (though their estimates are not exclusively of low skill workers).

worker has a higher realized productivity in the United States than in Bolivia—due to the institutions, technologies, and other complementary inputs available in the United States—and therefore would earn higher wages in the United States. Any policy that limits the realized productivity of that worker, by limiting access to the United States and its complementary inputs, obliges that person to accept a lower wage than he or she could potentially realize. This is wage discrimination based on a socially-constructed trait, workers’ country of their birth; economically its welfare impact on the worker is indistinguishable from forcing one of two workers of equal realized productivity to accept a lower wage because they were say, a woman or African-American. While in common usage “discrimination” and “prejudice” have become synonymous, no American need feel toward that worker any distaste or prejudice, racial or otherwise, in order to produce this result.<sup>47</sup>

In the United States, the wages of males to those of observably identical females who are in the workforce is about 1.3.<sup>48</sup> In our sample of 42 developing countries the median domestic ratio of wages of males to observably equivalent females is 1.4 (Madagascar), the large magnitude of which explains the justifiable widespread concern about sex and gender discrimination. But for several of the countries we consider in the preceding section, estimated border-induced wage discrimination vastly exceeds the estimate of domestic gender discrimination.

The literature contains several regressions exploring wage gaps between workers that are observably identical except for their ethnicity, within spatially integrated labor markets. These have included measures of the gaps for black workers in the United States, scheduled castes/scheduled tribes in India, indigenous Malaysians, Indigenous

---

<sup>47</sup> Some might question the use of the term “discrimination” for international differences. The foundational model of labor market discrimination by Becker (1971) (p. 35) actually begins—for conceptual clarity—as a model of labor migration from a black ‘country’ to a white ‘country’, and only later comes to represent racial discrimination in a single spatially integrated market by analogy. Becker defines wage discrimination as cross-group differences in wages unrelated to worker productivity, which is precisely what we seek here. He notes that “treating discrimination as a problem in trade and migration is far from artificial, since they are closely and profoundly related”. Arguments that migration restrictions constitute employment discrimination have been advanced in legal theory by Chang (2003) and in moral philosophy by Carens (1987).

<sup>48</sup> Altonji and Blank (1999) (Table 1, column 2) show that ln mean hourly wage for a full-year, full-time employed white male in 1995 is  $\ln(17.97) = 2.89$ . The coefficient on female in Table 4, column 6 is  $-0.241$ , suggesting the ln mean hourly wage for an otherwise observably identical woman is 14.12, and  $17.97/14.12 = 1.27$

Tanzanians, and indigenous Bolivians.<sup>49</sup> All of these reveal that wages in the base group exceed wages in the group discriminated against by a factor of 1.1–1.9 unexplained by characteristics. For example, the ratio of a white man’s wages to those of an observably identical black man in the United States was about 1.6 in 1939 (by 1995 it had fallen to about 1.1).<sup>50</sup>

The lesson of these numbers is that the wage discrimination created by international borders is, for large numbers of developing countries, at least as large as any current form of wage discrimination against socially disfavored groups within spatially integrated labor markets. For many of the 42 developing countries we investigate, it is much larger. For several countries—including Nigeria, Haiti, Egypt, and Ghana—the US border effect on the wages of equal intrinsic productivity workers is greater than *any* form of wage discrimination (gender, race, or ethnicity) that has *ever* been measured.<sup>51</sup>

---

<sup>49</sup> Banerjee and Knight (1991) report that in survey data collected in Delhi from October 1975 to April 1976, observably identical Indians who were not members of Scheduled Castes/Scheduled Tribes (SC/ST) earned 10.9% more than members of SC/STs (p. 185). Knight and Sabot (1991) find that in a survey of 1,000 urban Tanzanian firms in 1971, people who were not indigenous Africans (almost all of Asian ancestry) earned 87% more than observably identical indigenous Africans (p. 65). The Mincer wage regressions of Patrinos and Psacharopoulos (1993) (p. 305) indicate that a 28% wage deficit for indigenous workers relative to other workers is unexplained by differences in observable characteristics, thus  $1/(1-0.28) = 1.39$ . Schafgans (1998) (p. 483) finds a maximum differential in wages between Chinese and Malay males unexplained by observables of 27%, thus  $1/(1-0.27) = 1.37$ .

<sup>50</sup> Altonji and Blank (1999) (Table 1, column 2) show that ln mean hourly wage for a full-year, full-time employed white male in 1995 is  $\ln(17.97) = 2.89$ . The coefficient on black is  $-0.067$ , suggesting that ln mean hourly wage for an otherwise observably identical black male is 16.81, and  $17.97/16.81 = 1.07$ . Chandra (2000) documents the substantial narrowing of the black-white wage gap between 1940 and 1990, and Heckman, Lyons, and Todd (2000) suggest that some substantial portion of the narrowing during this period remains unexplained by changes in observable characteristics other than race. Sundstrom (2007) (p. 412) conducts Mincer regressions using the 1940 US census and finds a black-white wage gap of 36% unexplained by observable characteristics in 1939 wages.

<sup>51</sup> It is not clear that historical wage discrimination against African Americans, even at its most egregious—which is to say, when African Americans were routinely and forcibly held as property by other Americans—was ever as large as the wage discrimination caused by today’s limitation of movement between the US and the poorest countries. Wage discrimination against slaves (not the full effect of slavery, but exclusively the earnings effect) can be plausibly measured by comparing slave rental rates—reflective of slaves’ productivity indicator in a competitive rental market—to the cash value of owners’ maintenance costs. Though surviving data are controversial and inexact, some estimates suggest that US slaves’ productivity was roughly four times the cash value of food, shelter, clothing, and medical care that they received. This is well below today’s degree of wage discrimination produced by international borders against nationals of many development countries. Fogel and Engerman (1974) (II.159), for example, estimate \$60.62 as the cash value of average maintenance and compensation for adult male enslaved field hands on large plantations, and the same authors (II.73) present rental rates for the Lower South of \$143 in 1841-45, \$168 in 1846-50, \$167 in 1851-55, and \$196.5 in 1856-60; the simple average of these four figures is \$168.6. Note that these rental rates are net of maintenance costs (II.75), so the average rate inclusive of maintenance is \$229.2. In a competitive rental market, this would suggest that enslaved workers were producing 3.8 times what was spent to sustain them.

### 5.3 *Comparison with other antipoverty interventions for the marginal worker*

There is a sense of unease with the widening global inequalities (in absolute, and for some countries, relative terms) and a sense that “globalization” should be made more “equitable” by addressing the concerns of poor, a sense which has motivated social movements such as debt relief, “fair trade” movements, attempts to make the Doha round of WTO negotiations a “development” round, “anti-sweatshop” activism, the promotion of micro-credit, the evaluation of social programs, and so on. However, it should be obvious that if the marginal welfare gains from relaxing border constraints are larger. We have shown that these gains are on the order of P\$10,000 for workers from countries whose annual income is on the order of P\$5,000 per capita—the average for the developing world in aggregate. Movement must, therefore, reduce poverty more reliably and immediately for the marginal household than anything *in situ* antipoverty intervention under consideration.

Consider the documented wage gains due to microcredit provision, international labor rights campaigns, schooling incentives, and medical interventions. The pioneering (if controversial) estimate by Pitt and Khandker (1998) of the net return on microloans to Bangladeshi women is 18%, which means that the intervention multiplies household income by a factor of 1.14.<sup>52</sup> Harrison and Scorse (2004) (p. 21 and Table 1B) find that international anti-sweatshop campaigns against textile, footwear, and apparel plants in Indonesia in the early 1990s caused higher wages for workers at foreign-owned and export-focused plants by a factor of 1.20 to 1.25. In a typical country from our sample, the simple Mincer coefficient on years of schooling is 0.061,<sup>53</sup> meaning that three additional years of schooling somehow costlessly induced by an ideal incentive program would raise wages by a factor of around 1.2. Finally, Kremer and Miguel (2004) (p. 206),

---

<sup>52</sup> Their estimate of the return to males is 11%, but we use the higher figure for females to be conservative. Average annual female borrowing is Tk3415, or US\$361 at PPP using the average PPP conversion factor from World Bank (2007) over the relevant period (1986-1992) of 9.47. The resulting increase in household income is thus Tk615. Since annual income is Tk4372 (their Table A2), earnings rise by a factor of  $(4372+615)/4372 = 1.14$ .

<sup>53</sup> This is not the “return” to schooling, as Heckman, Lochner, and Todd (2006) stress; it does not account for costs of the education investment, most notably the opportunity cost of time. Strictly speaking, by the assumptions underlying the Mincer model the marginal return to schooling is zero, since individuals acquire education until the benefit equals the opportunity cost.

find that medicating children to eliminate hookworm, roundworm, and other helminthic parasites is “by far the most cost-effective method of improving school participation among a series of educational interventions” attempted in Busia, Kenya that underwent randomized evaluations. But the lifetime increment in wages thereby caused is 1.01.<sup>54</sup>

These values suggest that for a large number of countries, there is no rigorously evaluated *in situ* antipoverty intervention in the development economics literature capable of causing wage gains for the marginal household that come anywhere near those accruing from movement across borders. It is revelatory to calculate how long a worker would have to be allowed to work at US wages, for a willing employer as their wages are justified by their productivity in the US, in order to match the gains from other, widely popular, anti-poverty interventions. We have previously calculated that the net present value of *lifetime* access to micro-credit (again, under generous assumptions on returns) is the equivalent of the wage differential of a Bangladeshi man for *one month* of work in the US (Clemens, Montenegro and Pritchett, 2008).

This of course does not suggest that *in situ* interventions are not worth carrying out. Any intervention with a positive net present value is worth carrying out, and there is no reason not to do several in parallel. Furthermore, the general equilibrium effects of very large movements of people could differ from the marginal effects we measure, just as the general equilibrium effects of *in situ* antipoverty programs can differ enormously from the marginal effects measured by a field trial. That said, the marginal wage effects of movement can greatly exceed the marginal wage effects of *in situ* policies, and deserve a larger role in the discussion of development and antipoverty initiatives.

---

<sup>54</sup> The Mincer coefficient on years of schooling is 0.07, and the program causes an increase of 0.14 years in averages schooling.



## 6 Conclusion

The combination of wage surveys around the world with the US Census allows us to estimate cross-border wage differences for observably equivalent workers between the US and 42 developing countries. The median wage gap for a male, unskilled (9 years of schooling), 35 year-old, urban formal sector worker born and educated in a developing country is P\$15,400 per year at purchasing power parity. This figure understates the gain in household welfare caused by crossing the border if any of the wage increment is spent in the country of origin—as remittances or repatriated savings—where prices are lower and a dollar is worth more than a PPP dollar. To be conservative, we do not make any corresponding adjustment in our estimates.

On the other hand, the same estimate of the wage gain tends to overstate the true gains to a potential migrant both because of selection of migrants and because of welfare losses from moving. We offer a combination of new evidence on selection (in 9 of the 42 countries with new data) and wage differentials in legally integrated markets to suggest it would be conservative to scale back the wage ratios of observational equivalent workers by 1.5 to produce an estimate of the wage gains to a worker of equal intrinsic productivity who is willing to move.

We began this paper seeking to connect two empirical facts: the worldwide mesh of restrictions on labor movement, and large global differences in income. A conservative estimate of the welfare gain to a moderately skilled worker in the median country of our sample moving to the US is P\$10,000 per worker, per year, roughly double income per capita of the developing world in aggregate. We give evidence that gaps like these can only arise from policy barriers to labor movement. This means that those barriers create one of the largest distortions in any global market, create the largest form of wage discrimination observed in today's world, and create what is apparently the largest antipoverty intervention available for people from poor countries. The economic consequences of such barriers are clearly colossal, but remain largely unexplored.

## References

- Akee, Randall K. Q. 2007a. "Errors in Self-Reported Earnings: The Role of Previous Earnings Volatility." Bonn, Germany: Institute for the Study of Labor (IZA) Working Paper Series 3263.
- 2007b. "Who Leaves and Who Returns? Deciphering Immigrant Self-Selection from a Developing Country." Bonn, Germany: Institute for the Study of Labor Working Paper Series 3268.
- Altonji, Joseph, and Rebecca M. Blank. 1999. "Race and Gender in the Labor Market." In Orley Ashenfelter and David Card, (eds.), *Handbook of Labor Economics*. Amsterdam, Netherlands: Elsevier.
- Amuedo-Dorantes, Catalina, Cynthia Bansak, and Susan Pozo (2005), "On the remitting patterns of immigrants: Evidence from Mexican survey data", *Economic Review* 90 (1): 37-58.
- Anderson, James E. and J. Peter Neary. 1994. "Measuring the Restrictiveness of Trade Policy". *World Bank Economic Review* 8 (2): 151-169
- Banerjee, Biswajit and J. B. Knight. 1991. "Job discrimination and untouchability". In Nancy Birdsall and Richard Sabot, (eds.), *Unfair Advantage: Labor market discrimination in developing countries*. Washington, DC: World Bank.
- Becker, Gary S. 1971. *The Economics of Discrimination, 2nd Edition*. Chicago, IL: University of Chicago Press.
- Berry, William D., Richard C. Fording, and Russell L. Hanson. 2000. "An Annual Cost of Living Index for the American States." *Journal of Politics* 60(2):550–67.
- Borjas, George J. (1987), "Self-Selection and the Earnings of Immigrants." *American Economic Review* 77(4):531-553.
- Bound, John, and Alan B. Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9(1):1–24.
- Bratsberg, Berndt, and Dek Terrell. 2002. "School Quality and Returns to Education of US Immigrants." *Economic Inquiry* 40(2):177–98.
- Carens, Joseph H. 1987. "Aliens and Citizens: The Case of Open Borders." *Review of Politics* 49(2):251–73.
- Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences." In Philippe Aghion and Steven Durlauf, (eds.), *Handbook of Economic Growth. Volume 1A*. Amsterdam, Netherlands: Elsevier.

- Caselli, Francesco and James Feyrer, 2008, "The Marginal Product of Capital," *The Quarterly Journal of Economics*, vol. 122(2), pages 535-568
- Chandra, Amitabh. 2000. "Labor-Market Dropouts and the Racial Wage Gap: 1940-1990." *American Economic Review* 90(2):333-8.
- Chang, Howard. 2003. "Immigration and the Workplace: Immigration Restrictions as Employment Discrimination." *Chicago-Kent Law Review* 78(1):291-330.
- Clemens, Michael A., Claudio Montenegro, and Lant Pritchett. 2008, "The place premium : wage differences for identical workers across the US border", Policy Research Working Paper 4671. Washington, DC: World Bank.
- Cortes, Kalena. 2004. "Are Refugees Different from Economic Migrants? Some Empirical Evidence on the Heterogeneity of Immigrant Groups in the United States." *Review of Economics and Statistics* 86(2):465-80.
- Deaton, Angus and Alan Heston. 2008. "Understanding PPPs and PPP-based national accounts". Cambridge, MA: National Bureau of Economic Research Working Paper Series 14499.
- Easterly, William. 2004. "Channels from Globalization to Inequality: Productivity World versus Factor World." In Susan M. Collins and Carol Graham, (eds.), *Brookings Trade Forum 2004*. Washington, DC: Brookings Institution.
- Easterly, William, and Sergio Rebelo. 1993. "Marginal Income Tax Rates and Economic Growth in Developing Countries." *European Economic Review* 37(2-3):409-17.
- Fernández-Huertas Moraga, Jesús. 2008. "New Evidence on Emigrant Selection." Barcelona: Unitat de Fonaments de l'Anàlisi Econòmica (UAB) and Institut d'Anàlisi Econòmica (CSIC) in its series UFAE and IAE Working Papers 742.08.
- Fogel, Robert W., and Stanley L. Engerman. 1974. *Time on the Cross: The Economics of American Slavery (first edition, Vols. I & II)*. Boston: Little, Brown & Co.
- Freeman, Richard B., and Remco H. Oostendorp. 2005. "Occupational Wages Around the World". Cambridge, MA, National Bureau of Economic Research.
- Gollin, Douglas. 2002. "Getting Income Shares Right." *Journal of Political Economy* 110(2):458-74.
- Gould, Eric and Omer Moav (2008), "When is 'Too Much' Inequality Not Enough? The Selection of Israeli Emigrants", London: Centre for Economic and Policy Research Discussion Paper Series 6955.
- Hall, Robert E., and Charles I. Jones. 1998. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" Cambridge, MA: National Bureau of Economic Research Working Paper Series 6564.

- , 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics* 114(1):83–116.
- Harrison, Ann, and Jason Scorse. 2004. "Moving Up or Moving Out? Anti-sweatshop Activists and Labor Market Outcomes." Cambridge, MA: National Bureau of Economic Research Working Paper Series 10492.
- Hatton, Timothy J., and Jeffrey G. Williamson. 2005. *Global Migration and the World Economy: Two Centuries of Policy and Performance*. Cambridge, MA: Massachusetts Institute of Technology.
- Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. "Earnings Equations and the Rate of Return: The Mincer Equation and Beyond." In Eric A. Hanushek and F. Welch, (eds.), *Handbook of the Economics of Education*. Amsterdam: Elsevier.
- Heckman, James J., Thomas M. Lyons, and Petra E. Todd. 2000. "Understanding Black-White Wage Differentials, 1960-1990." *American Economic Review* 90(2):344–9.
- Hendricks, Lutz. 2002. "How important is human capital for development?" *American Economic Review* 92(1):198-219.
- Heston, Alan, Robert Summers, and Bettina Aten. 2006. *Penn World Tables Version 6.2*. Philadelphia, PA: University of Pennsylvania.
- Hoefert, Andreas, and Simone Hofer, eds. 2007. *Price and Earnings: A Comparison of Purchasing Power Around the Globe*. Zurich, Switzerland: Union Bank of Switzerland AG, Wealth Management Research.
- Jaeger, David A. 1997. "Reconciling the Old and New Census Bureau Education Questions: Recommendations for Researchers." *Journal of Business and Economic Statistics* 15(3):300–9.
- Klein, Paul and Gustav Ventura, 2004, "Do Migration Restrictions Matter?" Unpublished manuscript.
- Knight, J. B. and Richard H. Sabot. 1991. "Labor Market Discrimination in a Poor Urban Economy". In Nancy Birdsall and Richard Sabot, (eds.), *Unfair Advantage: Labor market discrimination in developing countries*. Washington, DC: World Bank.
- Kremer, Michael, and Edward Miguel. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72(1):159–217.
- Lamont, Owen A., and Richard H. Thaler. 2003. "The Law of One Price in Financial Markets." *Journal of Economic Perspectives* 17(4):191–202.
- Lewis, W. Arthur. 1954. "Economic development with unlimited supplies of labor." *Manchester School of Economic and Social Studies*, 22(2):139-191.

- Massey, Douglas S. and René Zenteno (2000), "A validation of the ethnosurvey: The case of Mexico-US migration", *International Migration Review* 34 (3): 766-793.
- McCallum, John. 1995. "National borders matter: Canada-U.S. regional trade patterns." *American Economic Review* 85(3):615-623.
- McKenzie, David J., John Gibson, and Steven Stillman. 2009. "How Important is Selection? Experimental Versus Non-Experimental Measures of the Income Gains From Migration." *Journal of the European Economic Association*, forthcoming.
- Milanovic, Branko. 2008. "Where in the World are you? Assessing the Importance of Circumstance and Effort in a World of Different Mean Country Incomes and (Almost) no Migration." Washington, DC: World Bank Policy Research Working Paper Series 4493.
- Montenegro, Claudio E., and Maximilian L. Hirn. 2008. "A New Set of Disaggregated Labor Market Indicators Using Standardized Household Surveys from Around the World." World Bank. Washington, DC. Mimeo.
- Office of Personnel Management (2000), "Report on 1998 Surveys Used to Determine Cost-of-Living Allowance in Non-Foreign Areas", *Federal Register* 65 (137): 44103-44173.
- O'Rourke, Kevin H., and Jeffrey Williamson. 2000. *Globalization and History: The Evolution of a Nineteenth-Century Atlantic Economy*. Cambridge, MA: MIT Press.
- Patrinos, Harry Anthony, and George Psacharopoulos. 1993. "The Cost of Being Indigenous in Bolivia: An Empirical Analysis of Educational Attainments and Outcomes." *Bulletin of Latin American Research* 12(3):293-309.
- Pitt, Mark M., and Shahidur R. Khandker. 1998. "The Impact of Group-based Credit Programs on Poor Households in Bangladesh: Does the Gender of Participants Matter?" *Journal of Political Economy* 106(5):958-96.
- Rama, Martin, and Raquel Artecona. 2002. "A Data Base of Labor Market Indicators across Countries." World Bank. Washington, DC. Mimeo.
- Rosenzweig, Mark. 2006. "Global wage differences and international student flows." In Susan M. Collins and Carol Graham, eds. *Brookings Trade Forum 2006: Global Labor Markets?* Washington, DC: The Brookings Institution.
- Roy, Andrew D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3(2):135-46.
- Schafgans, Marcia M. A. 1998. "Ethnic Wage Differences in Malaysia: Parametric and Semiparametric Estimation of the Chinese-Malay Wage Gap." *Journal of Applied Economics* 13(5):481-504.

- Semyonov, Moshe and Anastasia Gorodzeisky (2005), "Labor migration, remittances, and household income: A comparison between Filipino and Filipina overseas workers," *International Migration Review* 39 (1): 45-68.
- Shaw, William. 2007. *Migration in Africa: A Review of the Economic Literature on International Migration in 10 Countries*. Development Prospects Group. Washington, DC: The World Bank.
- Sundstrom, William A. 2007. "The Geography of Wage Discrimination in the Pre-Civil Rights South." *Journal of Economic History* 67(2):410-44.
- Vollrath, Dietrich. Forthcoming. "How Important are Dual Economy Effects for Aggregate Productivity?" *Journal of Development Economics*.
- Walmsley, Terrie, and L. Alain Winters. 2005. "Relaxing the Restrictions on the Temporary Movement of Natural Persons: A Simulation Analysis." *Journal of Economic Integration* 20(4):688-726.
- Williamson, Jeffrey G. 1995. "The Evolution of Global Labor Markets Since 1830: Background Evidence and Hypotheses," *Explorations in Economic History* 32 (2): 141-96.
- Williamson, Jeffrey G. 1997. "Globalization and Inequality, Past and Present." *World Bank Research Observer* 12 (2): 117-35.
- World Bank. 1995. *World Development Report 1995*. Washington, DC: World Bank.
- . 2005. *Global Economic Prospects: Trade, Regionalism and Development*. Washington, DC: World Bank.
- . 2007. *World Development Indicators 2007*. Washington, DC: World Bank.

**Appendix Table A1: Sample sizes for regressions underlying Table 1**

<b>Birth Residence Arrival</b>	<b>Total</b>	<b>US US —</b>	<b>Foreign US &lt; Age 20</b>	<b>Foreign US ≥ Age 20</b>	<b>Foreign Foreign —</b>
Argentina	475731	454503	1088	1969	18171
Bangladesh	459990	454503	412	1611	3464
Belize	456241	454503	476	624	638
Bolivia	459128	454503	502	953	3170
Brazil	565725	454503	1422	3740	106060
Cambodia	464296	454503	1687	1743	6363
Cameroon	457004	454503	55	282	2164
Chile	514339	454503	675	1308	57853
Colombia	491547	454503	4173	8808	24063
Costa Rica	465756	454503	766	1184	9303
Dominican Republic	473726	454503	6528	9679	3016
Ecuador	477768	454503	2722	5028	15515
Egypt	461275	454503	708	2088	3976
Ethiopia	478235	454503	515	1457	21760
Ghana	459889	454503	309	1593	3484
Guatemala	470657	454503	5115	8465	2574
Guyana	461272	454503	1988	3701	1080
Haiti	466569	454503	3522	8026	518
India	564435	454503	4897	22593	82442
Indonesia	565564	454503	717	1048	109296
Jamaica	472060	454503	6099	9375	2083
Jordan	466559	454503	465	604	10987
Mexico	718994	454503	116752	129862	17877
Morocco	459465	454503	349	709	3904
Nepal	455887	454503	77	251	1056
Nicaragua	464411	454503	2712	3506	3690
Nigeria	460931	454503	697	3091	2640
Pakistan	471846	454503	1303	3628	12412
Panama	470010	454503	2097	1755	11655
Paraguay	460959	454503	77	175	6204
Peru	475229	454503	2207	5472	13047
Philippines	532727	454503	13696	30630	33898
Sierra Leone	455581	454503	109	503	466
South Africa	474877	454503	531	1188	18655
Sri Lanka	471740	454503	134	603	16500
Thailand	486832	454503	1984	2256	28089
Turkey	528367	454503	772	1332	71760
Uganda	457497	454503	71	290	2633
Uruguay	473507	454503	238	424	18342
Venezuela	487625	454503	1041	1623	30458
Vietnam	501954	454503	10856	16712	19883
Yemen	465329	454503	143	194	10489

## APPENDIX: DATA

### A1 Sources

Survey data on wages and other worker characteristics from 42 developing countries were compiled by Indermit S. Gill and his team at the World Bank. Appendix Table A3 lists the original sources and size of each sample, as well as reproducing the exact text of the wage question from each survey. A detailed description of the database can be found in Montenegro and Hirn (2008).

In three surveys (India, Turkey and the US), the respondent's education level is listed as achievement categories rather than as years of schooling. We translate these categories into years of schooling according to the information available in the surveys. In the particular case of the US we use the following concordance: 0 years if "less than 1st grade"; 3 years if "1st through 4th grade"; 5.6666 years if "5th or 6th"; 7.6666 years if "7th or 8th"; 9, 10, 11, or 12 have separate categories; 12 years if "high school equivalent"; 13.5 years if "some college but no degree"; 14 years if "associate degree" or equivalent; 16 years if "bachelor's degree"; 18 years if "master's degree"; 19 years if "professional degree"; 20 years if "doctoral degree". This is a compromise blend of the methods used in Bratsberg and Terrell (2002) and in Jaeger (1997).

All data except the cost of living index used in the wage regressions for US states and Puerto Rico come from the United States Public Use Microdata Sample (5%) of the 2000 census. US state and Puerto Rico cost of living index comes from the revised 2004 version of the Berry-Fording-Hanson (BFH) state cost of living index (described in Berry, Fording, and Hanson (2000)), which excludes Alaska, Hawaii, and the District of Columbia. In the BFH index for 1999, 1 is the purchasing power of \$1 in the median US state. For Puerto Rico we use 0.86, which is the PPP conversion factor for 1999 from the Penn World Table 6.2 (Heston, Summers, and Aten (2006)).

### A2 Quality

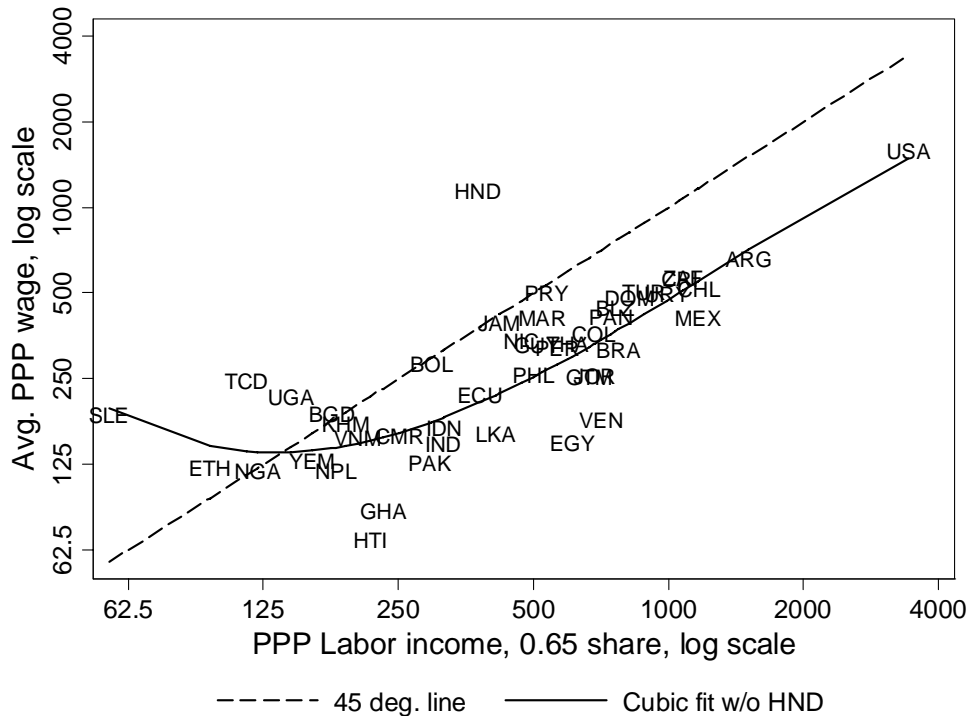
An important question we do not take up in the text is the degree to which the surveys are representative of the wage sector and of the country—though all were designed to be. One way to check the representativeness of the wage surveys is to compare national accounts estimates of labor income per worker at PPP in each country to the average wage from the surveys we use. There is no reason to expect these to be equal—most importantly because the wage data we use do not include self-employed people and therefore do not include large portions of the informal sector, and even informal-sector wage workers can be harder to sample than formal-sector workers. But enormous differences between the two could signal problems in the representativeness of the survey data.

Appendix Figure 1 plots this comparison. Labor income per worker is calculated by the method of Gollin (2002), under the assumption that a 0.65 share of GDP accrues to labor.



The dotted line shows a 45-degree line and the solid line shows a cubic least-squares fit to the data including a dummy for Honduras ( $R^2 = 0.756$ ; or  $R^2 = 0.607$  without the Honduras dummy). Large amounts of self-employment would tend to push countries down and to the right; large amounts of low-wage informal sector work would tend to push countries up and to the left.

**Appendix Figure A1: Comparison of labor income per worker and survey-based wages**



Line shows cubic regression fit of  $\ln(\text{wage})$  on  $\ln(\text{labor income})$ , its square, and its cube, with a dummy for Honduras.

We draw three lessons from Figure 1. First, the agreement is in general very good. Average wage is typically some reasonable fraction of average labor income, varying across countries as would be expected given different relative sizes of the informal sector and the self-employed sector. Second, formal-sector wages are clearly not representative of typical worker earnings in the poorest countries with very large informal sectors (e.g. Sierra Leone, Chad, Nigeria, Bangladesh, Ethiopia, Uganda). This is to be expected given that earnings gaps between the formal and informal sector are highest in the poorest countries (Vollrath (Forthcoming)). Third, the Honduras survey appears anomalous and we drop Honduras from subsequent reported analysis (since our preferred results are from bi-lateral regressions this has no consequence for other countries' results). Overall, this analysis highlights the fact that all of the estimates to follow can only be interpreted as applying to a worker moving across the formal wage-labor sector. If indeed the formal-informal gap is much larger in poor countries, this *underestimates* the wage gains for a typical worker.

Appendix Table A2: Household survey data sources

Country	Year	Survey	Survey agency	Sample	Wage question
Argentina	2001	Encuesta Permanente de Hogares	Instituto Nacional de Estadísticas y Censos (INDEC)	19,706	Cuanto cobró por ese mes por esos conceptos? (Monto total de sueldos/jornales, salario familiar, horas extras, otras bonificaciones habituales y tickets vales o similares)
Bangladesh	2000	Household Income Expenditure Survey	Bureau of Statistics	3,517	What is your total net take-home monthly cash remuneration after all deductions at source?
Belize	1995	Survey of Living Conditions	Central Statistical Office	783	What is your gross monthly income?
Bolivia	2002	Encuesta de Hogares	Instituto Nacional de Estadísticas	3,244	Cuál es el salario líquido de su trabajo en horario normal?
Brazil	2005	Pesquisa Nacional per Amostra de Domicílios	Instituto Brasileiro de Geografia e Estatística	107,955	Qual era o rendimento mensal que você ganhava normalmente em setembro de 2003, nesse trabalho?
Cambodia	2004	Household Socio-Economic Survey	National Institute of Statistics	8,578	How much did you earn in salary/wages from this activity last month?
Cameroon	2001	Enquête Camerounaise Auprès de Ménages	Direction de la Statistique et de la Comptabilité Nationale	5,098	A quel montant estimez vous la totalité des revenus issus de votre emploi principal le mois dernier?
Chile	2003	Encuesta de Caracterización Socio-económica Nacional	Ministerio de Planificación	59,532	En el mes pasado, cuál fue su ingreso o remuneración líquida en su ocupación principal?
Colombia	2000	Encuesta Continua de Hogares	Departamento Administrativo Nacional de Estadística	27,996	Cuanto ganó el mes pasado en este empleo? (incluya propinas y comisiones y excluya viáticos y pagos en especie)
Costa Rica	2001	Encuesta de Hogares de Propósitos Múltiples	Instituto Nacional de Estadísticas y Censos	12,503	En su ocupación principal, cuál fue el ingreso efectivamente percibido por concepto de sueldo, salario, jornal, propinas, horas extras, en el último periodo de pago (semana, quincena o mes)?
Dominican Republic	1997	Encuesta Nacional de Fuerza de Trabajo	Departamento de Cuentas Nacionales y Estadísticas Económicas del Banco Central	3,056	Cuánto le pagan o gana usted y cada qué tiempo en ese trabajo?
Ecuador	2004	Encuesta de Empleo, Desempleo y Subempleo	Instituto Nacional de Estadísticas y Censos	17,576	En su ocupación cuánto dinero líquido recibió por concepto de sueldo o salario u otros ingresos en el mes de marzo?
Egypt	1998	Labor Market Survey	Central Agency for Public Mobilization and Statistics	4,776	What is the net amount received in basic net wage?
Ethiopia	2005	National Labour Force	Central Statistical Authority	22,568	What was the amount paid in your main occupation during the last period?
Ghana	1991	Living Standards Surveys Round Three	Statistical Office	5,749	What is the amount of money you will receive for this work?
Guatemala	2002	Encuesta Nacional Sobre Condiciones de Vida	Instituto Nacional de Estadísticas	2,584	Cuál es el último ingreso neto o ganancia que recibió?
Guyana	1992	Living Standards Measurement Survey	Bureau of Statistics	1,266	What is your cash income from paid employment (BASIC wage or salary)?
Haiti	2001	Les Conditions de Vie en Haïti	Institut Haïtien de Statistique et d'Informatique	1,220	What is your wage, salary, commission payments, bonuses or other cash income (including overtime) from employer?
India	1999	Socio-economic Survey	National Sample Survey Organization	94,306	What are the wage and salary earnings (received or receivable) for the work done during the week?
Indonesia	2002	Survei Sosial Ekonomi Nasional	Badan Pusat Statistik	129,279	How much is the wage/net salary received in a month of main work?
Jamaica	2002	Jamaica Survey of Living Conditions	Statistical Institute of Jamaica	3,723	What is the value of all income received in cash or in kind during the past 12 months?
Jordan	2002	Household Income Expenditure Survey	Household Surveys Directorate	12,824	What is the total income from employment?
Mexico	2002	Encuesta Nacional de Ingresos y Gastos de los Hogares	Instituto Nacional de Estadística, Geografía e Informática	18,064	Cuánto recibió el mes pasado por sueldos, salarios y jornales en el mes pasado? (declare su ingreso bruto)
Morocco	1998	Enquête Nationale sur les Niveaux de Vie des Ménages	Secretariat d'État à la Population, Direction de la Statistique	4,043	Quel a été votre salaire en espèce dans votre travail ?
Nepal	2003	Living Standards Survey II	Central Bureau of Statistics	2,216	How much did you get in cash per day for this job?

Appendix Table A3, continued: Household survey data sources

Nicaragua	2001	Encuesta Nacional de Hogares Sobre Medición de Nivel de Vida	Instituto Nacional de Estadísticas y Censos	3,757	Cuál es el último ingreso neto que tuvo usted?
Nigeria	2003	Living Standards Surveys	Federal Office of Statistics	3,084	What is the amount of money you received or you will receive for this work?
Pakistan	2001	Integrated Household Survey	Federal Bureau of Statistics	13,186	How much is your take-home pay, including bonuses or cash allowances?
Panama	2003	Encuesta de Hogares	Dirección de Estadística y Censo	14,392	Cuál es salario o ingreso mensual en su trabajo? (si es empleado investigue sueldos y salarios brutos—sin deducir impuestos ni contribuciones al seguro social)
Paraguay	2001	Encuesta Permanente de Hogares	Dirección General de Estadísticas, Encuestas y Censos	6,254	Cuál fue el monto del último pago neto o líquido que recibió (incluyen descuentos por préstamos, asociaciones, etc.)? Si no le han pagado todavía, cuánto espera que le paguen y que periodo de tiempo incluye este pago?
Peru	2002	Encuesta Nacional de Hogares	Instituto Nacional de Estadísticas e Informática	13,367	Cuál fue ingreso total en el pago anterior incluyendo horas extras, bonificaciones, pago por concepto de refrigerio, movilidad, comisiones, etc.?
Philippines	2002	Annual Poverty Indicators Survey	National Statistics Office	34,626	Total Income, salary/wages from employment
Sierra Leone	2003	Integrated Household Survey	Statistics Sierra Leone	565	What is the amount of money you received or you will receive?
South Africa	2000	Labour Force Survey	Statistics South Africa	21,707	What is your total salary/pay in your main job?
Sri Lanka	2002	Household Income and Expenditure Survey	Department of Census and Statistics	16,772	What is the wage/salary you received last calendar month?
Thailand	2002	Socio-economic Survey	National Statistical Office	28,258	Wage and salaries
Turkey	2005	Household Labour Force Survey	State Institute of Statistics	75,610	How much did you earn from main job activity during the last month?
Uganda	2002	Socio-economic Survey	Uganda Bureau of Statistics	3,204	How much do you earn per period? (effort should be taken to get the net salary after the deduction of taxes)
United States	1999	2000 Census Population and Housing (Public Use Microdata Sample)	US Census Bureau	1,124,253	Wages, salary, commissions, bonuses, or tips from all jobs. Report amount before deductions for taxes, bonds, dues or other items
Uruguay	1995	Encuesta Continua de Hogares	Instituto Nacional de Estadísticas	19,142	Cuánto ganó el mes pasado como empleado u obrero del sector público o privado?
Venezuela	2004	Encuesta de Hogares por Muestreo Nacional	Instituto Nacional de Estadísticas	34,569	En un mes normal cuánto es su ganancia neta?
Vietnam	2002	Household Living Standard Survey	General Statistical Office	19,920	In the past 12 months, how much did you receive from this work in money and in kind?
Yemen	2005	Household Budget Survey	Central Statistical Organization	10,583	How much was your last pay? (net of taxes and any other deduction)

Reported sample size corresponds the number of people in each survey who are wage earners and of age 15-65. Of the US sample, 500,319 were born in the US.