

## Technical appendix to “Income per natural: Measuring development for people rather than places”

Michael Clemens and Lant Pritchett  
July 2008

This note presents technical details of calculations and data sources in the main paper.

### Method

Table A1 presents summary statistics for the dataset underlying the model used to predict income per capita of the foreign-born in OECD countries. Table A2 presents the regression model estimated using US census data. The coefficients from the fourth column of Table A2 are used to predict income per capita of the foreign-born in each country of destination as a fraction of average income per capita in that country. Figure A1 presents an example of how the different coefficients estimated in Table A2, which uses data on the foreign-born in the US, are used to estimate the incomes of Mexican-born and Somalian-born people in Australia.

How much of the cross-country variance in income per natural is due to income gaps between sending and receiving countries  $\left(\frac{y_i^* - y_i}{y_i}\right)$ , and how much is due to variance in diaspora size across countries of origin ( $\theta_i^*$ )? For example, income per Pakistani natural goes up with the gap between incomes of those living in Pakistan those living in the UK, and also goes up with the number of Pakistanis living in the UK. Figure A2 plots the percent difference in natural population and origin-country resident population on the horizontal axis, and the percent difference in naturals’ income per capita in *OECD destination countries* and origin-country GDP per resident on the vertical axis. The dashes mark the 45° line  $\left(\theta_i^* = \left(\frac{y_i^* - y_i}{y_i}\right)\right)$ . Note that the percent differences are so large in some cases that both axes are on a  $\log_{10}$  scale.

The overwhelming determinant of the variance in income per natural is the income gaps. The vast majority of countries lie far to the upper left of the 45° line. Even

countries with a very high proportion of naturals abroad have income gaps much, much higher: for Jamaica  $\theta_i^* = 0.325$  and  $\left(\frac{y_i^* - y_i}{y_i}\right) = 3.31$ . Below the line are traditionally

sending countries whose domestic economy has improved—such as Ireland.

Interestingly, Mexico lies directly on the line: Household income per capita for Mexican-born in the US in 2000 was \$9,991, and Mexican GDP per resident at purchasing power parity is \$9,197. So Mexican income per natural would go up by roughly the same percentage due to a given percentage increase in the income gap or the same percentage increase in the number of Mexicans living in the US.

Table A3 shows the calculations based on Commander et al. (2004) of selection bias to our estimates using data on wages of Indian software workers in and out of India. It uses their wage figures for 1999, the year closest to the 2000 census data used in the rest of this paper, and gives the resulting estimates of the selection term in equation (2). Figure A3 summarizes this and the other examples of the calculation of selection bias.

### Data Sources

**Household income per capita of foreign born in the United States:** The data come from the Public Use Microdata Sample of the 2000 US Census. First, each person in the individuals file is matched to a household from the households file. Household income per capita for each individual is then calculated by dividing “Household Total Income in 1999” by “Number of Person Records Following This Housing Record”. All individuals born in the US are then dropped, and household income per capita is averaged across all foreign-born individuals separately by country of birth.

**Household income per capita of foreign born in Australia:** The data were prepared in custom extract from the master 2001 census file by the Information Consultancy Service of the Australian Bureau of Statistics, December 19, 2007. The extract consists of a crosstabulation of birthplace of individuals by average household income per person, and average household income per capita of individuals for persons enumerated at their usual residence on census night.

**GDP per capita at Purchasing Power Parity:** World Bank's *World Development Indicators 2007* and CIA *World Fact Book* (various years). Figures are for the year 2000 or closest available year. GDP per capita at PPP is assumed to be the same in Liberia as in Sierra Leone.

**Fraction of diaspora with tertiary education, and fraction with primary education or less:** Calculated from data in Dumont and Lemaître (2005). Includes people aged 15+ living in all OECD countries, at the time of the last census (circa 2000).

**Fraction of worldwide diaspora residing in each receiving country:** Data are from Parsons et al. (2007). The numerator is the number of people born in each sending country living in the receiving country in 2000, and the denominator is the number of people born in that sending country living anywhere outside the sending country in 2000.

**Fraction of OECD-resident diaspora residing in each receiving country:** Data are from Parsons et al. (2007). The numerator is the number of people born in each sending country living in each OECD receiving country in 2000, and the denominator is the number of people born in that sending country living anywhere in the OECD in 2000.

**Land area:** World Bank's *World Development Indicators 2007*. In square kilometers.

**Size of diaspora in each receiving country:** Parsons et al. (2007).

**Fraction of diaspora that is refugees:** Refugee populations in 2000, by origin and destination country, come from the United Nations High Commission for Refugees (UNHCR) Online Population Statistics Database, <http://www.unhcr.org/statistics>, downloaded July 31, 2007. The database defines refugees as “persons recognized as refugees under the 1951 Convention/1967 Protocol, the 1969 OAU Convention, in accordance with the UNHCR Statute, persons granted a humanitarian or comparable status and those granted temporary protection.”

**Population:** UN Population Division *World Population Prospects* database, 2006 revision, <http://esa.un.org/unpp>, downloaded July 31, 2007.

**Distance to country of birth, dummy for country of birth contiguous to destination country, dummy for common official language with destination country:** All are from the Centre des Études Prospectives et d'Informations Internationales (CEPII) bilateral distance database, <http://www.cepii.fr/anglaisgraph/bdd/distances.htm>, downloaded July 31, 2007.

**Dummy for oil-exporting Middle East:** Includes Saudi Arabia, Kuwait, Bahrain, Qatar, and United Arab Emirates.

**Poverty headcounts:** World Bank Povcalnet, download August 13, 2007:  
<http://iresearch.worldbank.org/PovcalNet/jsp/index.jsp>

Data sources for infant mortality per natural calculation

**Bilateral stocks of foreign-born** are from the Global Migrant Origin Database, Ver. 4, from the Development Research Center on Migration, Globalization and Poverty ([http://www.migrationdrc.org/research/typesofmigration/global\\_migrant\\_origin\\_database.html](http://www.migrationdrc.org/research/typesofmigration/global_migrant_origin_database.html), accessed 7 November 2007).

**Crude birth rates (CBR), population, and infant mortality rates (IMR)** for sending countries and host OECD countries are from the World Bank's *World Development Indicators 2007*. The CBR is recorded per 1,000 people, and IMR is recorded per 1,000 live births. For most countries, CBR and IMR data were available in 2000, and in the few countries where such data were missing, a linear interpolation was applied using the nearest available observations.

## References

- Commander, Simon, Rupa Chanda, Mari Kangasniemi, and L. Alan Winters (2004), "Must Skilled Migration Be a Brain Drain? Evidence from the Indian software industry," Discussion Paper 1422 (Bonn: Forschungsinstitut zur Zukunft der Arbeit [IZA]).
- Dumont, Jean-Christophe and Georges Lemaître (2005), "Counting Immigrants and Expatriates in OECD Countries: A New Perspective", OECD Social Employment and Migration Working Papers, No. 25 (Paris: OECD Publishing).
- Parsons, Christopher R., Ronald Skeldon, Terrie L. Walmsley, and L. Alan Winters (2007), "Quantifying international migration: A database of bilateral migrant stocks", in Çağlar Özden and Maurice Schiff, eds., *International Migration, Economic Development, and Policy* (Washington, DC: World Bank), pp. 17-58.

**Table A1: Summary statistics for the underlying dataset**

Variable	Obs	Mean	Std. Dev.	Min	Max
ln foreign-born income*	130	9.956466	.3497694	8.97702	10.59926
ln GDP/cap at PPP	4739	8.526976	1.169391	6.137727	10.56618
ln tertiary educ. frac.	3820	-1.26909	.6001354	-5.075174	0
ln primary educ. frac.	3769	-1.337871	.6853708	-4.582833	0
ln tertiary enrollment	4256	2.553927	1.3495	-1.609438	4.574711
ln world frac. in dest.	4588	-6.512936	2.693019	-14.11462	-.0826165
ln OECD frac. in dest.	4588	-5.570897	2.579269	-13.50135	0
ln land area	4739	10.98639	2.890235	.6678294	16.61165
ln size diaspora	4588	5.923727	3.107219	0	16.04947
Fraction refugees	4588	.0578732	.7223542	0	43.03552
ln dist. to ctry. birth	5152	8.673961	.8519143	2.951104	9.88258
Contiguous	5152	.0122283	.1099139	0	1
Common official lang.	5152	.1383929	.3453454	0	1
Oil-exporting Mid. East	5518	.0208409	.1428644	0	1
Latin Am./Carib.	5518	.1583907	.3651401	0	1
E. Europe/Central Asia	5518	.116709	.3211023	0	1
Mid. East/No. Africa	5518	.0873505	.2823736	0	1
Sub-Saharan Africa	5518	.1959043	.3969312	0	1
E. Asia/Pacific	5518	.141718	.3487923	0	1
South Asia	5518	.0333454	.1795532	0	1

\* “ln foreign-born income” refers to average ln Household income per person for those born in each sending country and residing in the United States in 2000 (one observation per country of birth). These 130 observations are used to fit the model; other observations are used on the right-hand side to make out-of-sample left-hand side predictions.

**Table A2: Correlates of household income per capita among the foreign-born in the United States**

Dependent variable: Average ln Household income per person for those born in each sending country and residing in the United States in 2000 (one observation per country of birth).

ln GDP per capita at PPP in country of birth	0.18 (7.93)	0.21 (6.14)	0.14 (5.53)	0.14 (5.60)
ln Fraction of diaspora with tertiary education			0.45 (6.24)	0.39 (6.33)
ln Fraction of diaspora with primary education or less			0.04 (0.57)	0.06 (1.08)
ln Gross tertiary enrollment in country of birth				0.04 (2.20)
ln Fraction of worldwide diaspora residing in US				0.11 (4.48)
ln Fraction of OECD-resident diaspora residing in US				-0.06 (2.44)
ln Land area of country of birth				0.03 (3.22)
ln Size of diaspora in US				-0.05 (3.07)
Fraction of US diaspora that is refugees				-0.44 (4.29)
ln Distance to country of birth				0.01 (0.14)
Dummy for country of birth contiguous to US				-0.10 (1.15)
Dummy for common official language with US				0.04 (1.34)
Dummy for oil-exporting Middle East				-0.47 (4.72)
Region dummy: Latin America & Caribbean		-0.37 (5.17)	-0.23 (4.65)	-0.18 (3.88)
Region dummy: Eastern Europe & Central Asia		-0.11 (1.46)	-0.22 (3.26)	-0.09 (1.35)
Region dummy: Middle East & North Africa		-0.09 (0.93)	-0.23 (2.69)	-0.02 (0.43)
Region dummy: Sub-Saharan Africa		0.06 (0.41)	-0.14 (1.45)	-0.01 (0.14)
Region dummy: East Asia & Pacific		-0.27 (2.88)	-0.27 (5.08)	-0.17 (2.39)
Region dummy: South Asia		0.07 (0.50)	-0.23 (1.96)	0.05 (0.43)
Constant	8.36 (40.91)	8.34 (24.49)	9.53 (32.32)	9.57 (26.32)
N	130	130	129	118
R <sup>2</sup>	0.326	0.536	0.803	0.893
Adjusted R <sup>2</sup>	0.321	0.509	0.788	0.872

Absolute value of robust t-statistics in parentheses. Omitted region dummy covers high-income OECD countries.

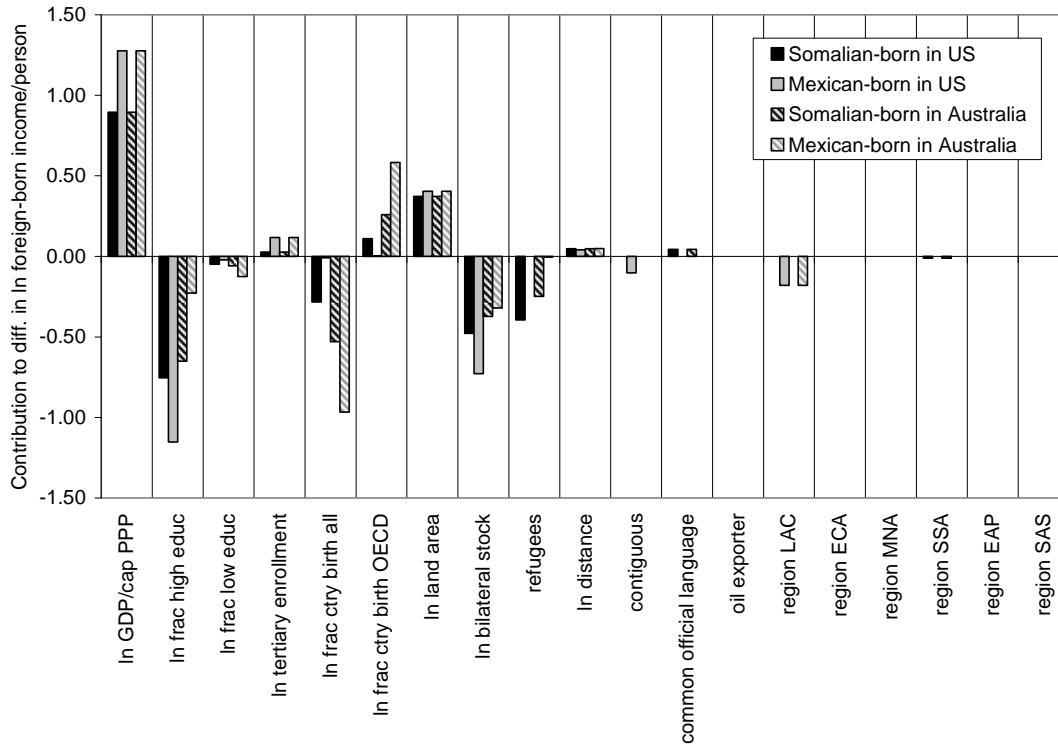
**Table A3: Counterfactual wages of Indian-born software industry workers in the US**

Workers	US wage, $y_i^*$	False counterfactual (GDP/cap.), $y_i$	True counterfactual, $\bar{y}_i^*$	Difference term $\frac{y_i^* / y_i - 1}{y_i^* / y_i - \bar{y}_i^* / y_i}$
Managers	83,762	2,311	20,391	1.29
Conceptualizers	72,291	2,311	17,178	1.27
Developers	64,448	2,311	11,916	1.18
Modifiers	42,068	2,311	9,071	1.20
Supporters	40,115	2,311	6,490	1.12

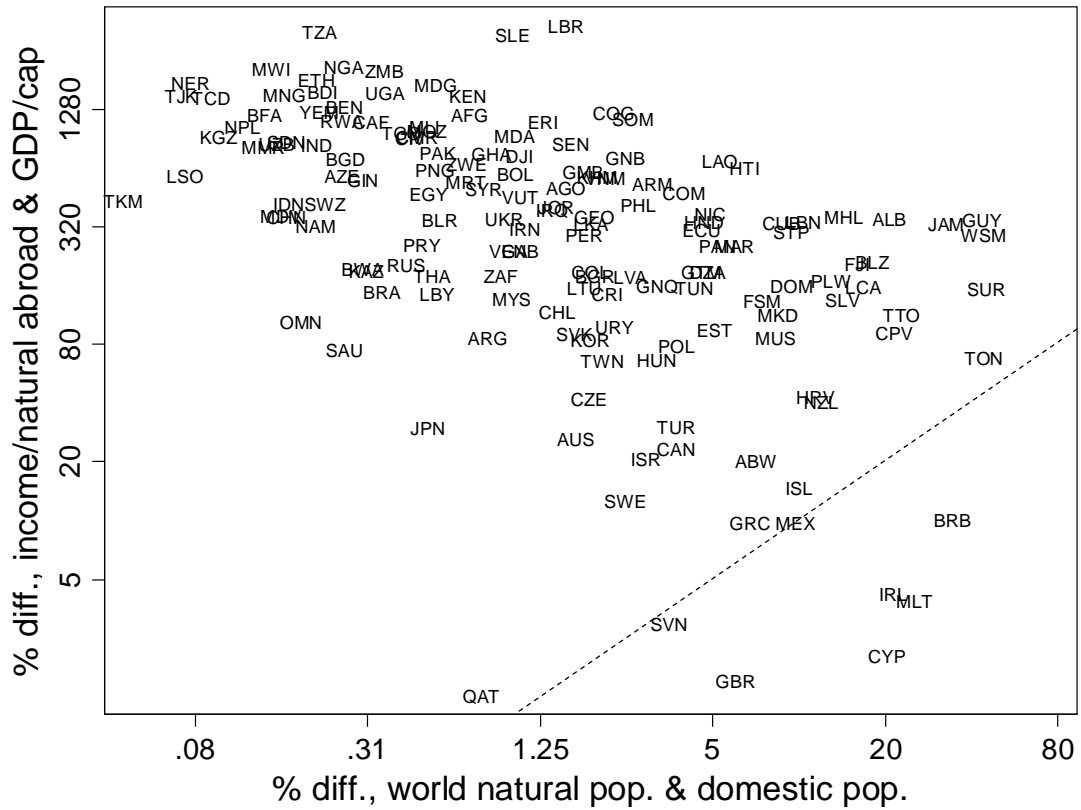
All dollar amounts are 1999 dollars per year at purchasing power parity. US wages and true counterfactual wages are from Commander et al. (2004), and GDP per capita is from the World Bank's *World Development Indicators 2007*.  $y_i^*$  is income per person in destination country,  $y_i$  is income per resident in origin country, and  $\bar{y}_i^*$  is counterfactual income per person if migrants had not left origin country.

**Figure A1:** Example estimation of foreign-born income/person for Somalia and Mexico

The vertical axis of the figure shows the product of the respective coefficient from Table 1, column 4, and the value of the respective variable for each migrant-sending and migrant-receiving country pair.

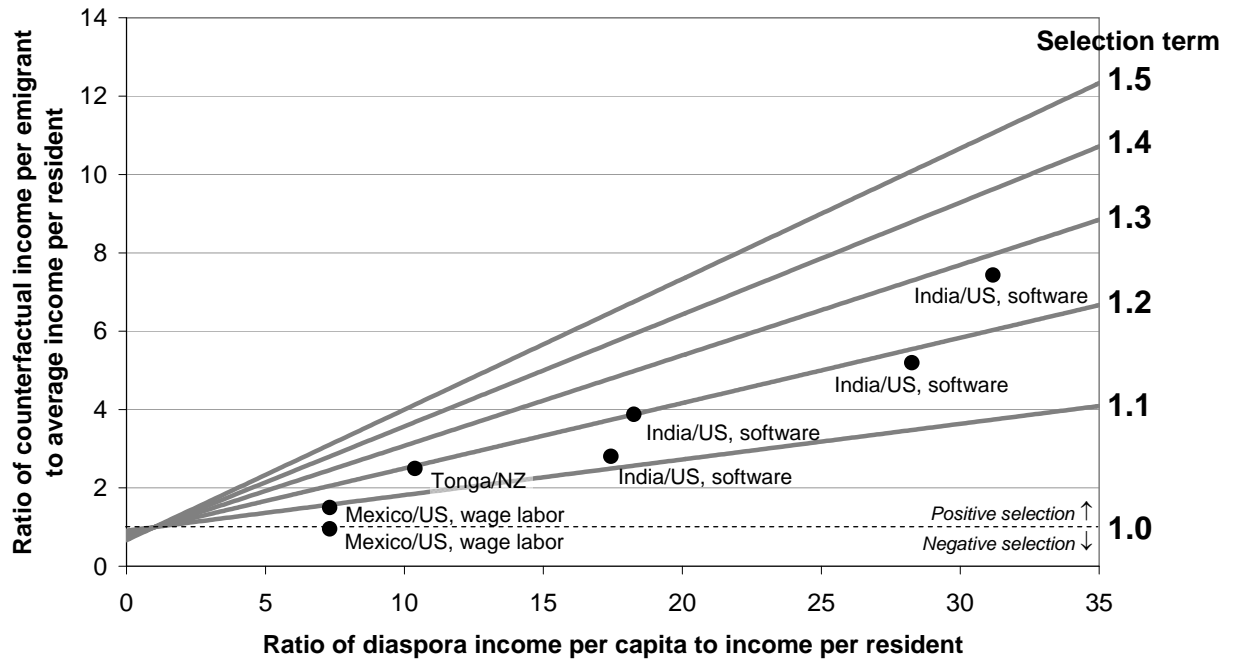


**Figure A2: Decomposition of cross-country variance in income per natural**



Dashes show 45° line. Both axes in log<sub>10</sub> scale. The horizontal axis shows the percent difference between total global population born in each country and the total population residing in each country. The vertical axis shows the percent difference between the average income of nationals of each country who reside abroad (exclusive of those who reside in their country of birth) and GDP per capita at PPP of the country.

**Figure A3: Size of the selection term in different scenarios**



Dots show estimates of the selection term calculated in the text of Section 3.2. Gray lines show the value of the selection term in equation (2) at different values of  $y_i^* / y_i$  (horizontal axis) and  $\bar{y}_i^* / y_i$  (vertical axis). Indian software industry top managers not shown.