

# Bubble Rap: Visualizing Poverty Dynamics

**Jonathan Haughton**

*Department of Economics, Suffolk University, Boston*

*Animated bubble graphs can be very useful in helping one to visualize the evolution of poverty. We discuss the issues involved in measuring global poverty, and then explain how to create these graphs in Excel, showing the links between poverty and GDP per capita for six regions of the world between 1981 and 2008. The animated graphs can provide insights that are difficult to gather from tables of numbers or static graphs.*

Keywords: *Poverty dynamics, bubble graphs*

## Introduction

In 2000, the United Nations adopted the Millennium Development Goals. The first, and most important, of the eight goals was a commitment to halve the percentage of the world population living in extreme poverty between 1990 and 2015.

Is the world on track towards meeting this key goal? Which regions are witnessing the greatest reductions in poverty? And what parts of the world are still poor? It would be very helpful to be able to visualize these effects, and in this short paper we explain how to do so in Excel using dynamic bubble graphs, a striking visual technique popularized by Hans Rosling.

We begin by discussing how to measure extreme poverty, and then introduce a series of increasingly complex graphs, which we may use to present the information efficiently.

## Measuring Poverty

Our interest here is in measuring poverty for major regions of the world, over time, although we could disaggregate by individual countries if we were so inclined. Following the World Bank (2000), we define poverty as a

“pronounced deprivation in well-being”. In order to make international comparisons of poverty, we need to make this definition operational. The most widely-recognized approach is that taken by Chen and Ravallion (2001, 2004, 2008), which establishes a world poverty line, translates it into domestic currencies using purchasing power parity, and then infers the poverty rate using information derived from household surveys. Some further discussion on each step is in order; for more details, see Haughton and Khandker (2009, chapter 10).

The process starts by establishing a poverty line, below which a person is considered to be poor. Chen and Ravallion (2008), guided by the official poverty lines in the poorest 15 countries for which they have data, argue that a good place to begin is with a poverty line of US\$1.25 per person per day (in 2005 prices). This is the successor to the original “dollar a day” poverty line that was used by the World Bank prior to 2008. The choice of a poverty line is somewhat arbitrary, even if it is motivated by an attempt to measure the cost of basic needs, and so in most analyses of poverty it is sensible to check the sensitivity of the results to other poverty lines such as, for example, US\$2.00 per person per day.

Given a poverty line, we now need to compute the equivalent value in local currency units. For this it is not appropriate to use the market or official exchange rate, because this is a poor guide to purchasing power: consumer prices are systematically higher in richer than in poorer countries (the Penn effect, often traced primarily to the Balassa-Samuelson effect; Balassa 1964, Samuelson 1964). Thus, for instance, a dollar in the U.S. buys less than a dollar converted into 52 rupees (the exchange rate in January 2012) and spent in India. A solution is to convert dollars into other currencies in such a way as to reflect purchasing power parity (PPP); with such an adjustment, perhaps the US dollar would be worth a hundred rupees, for example. This requires an enormous effort to gather data on prices in countries around the world; the International Comparison Project, run by the United Nations and the University of Pennsylvania, expanded and revised the computations of PPP for 2005, and these are the conversion rates that the World Bank used to turn the US\$1.25 poverty line into local-currency equivalents for 2005 (Penn World Tables ver. 6.3: see [http://pwt.econ.upenn.edu/php\\_site/pwt\\_index.php](http://pwt.econ.upenn.edu/php_site/pwt_index.php)).

Having translated the universal poverty line into local currencies for 2005, Chen and Ravallion (2008) then use local consumer price indexes to establish the poverty lines for other years. Then, for the years in which the information was collected, they use measures of consumption or income per capita from household survey data – or, more frequently, published tabulated data to which a distributional curve may be fitted – to compute the poverty rates for as many countries and years as there are available surveys. The national measures of poverty are then established, by interpolation if necessary, for every third year, starting in 1981, and aggregated to the global level.

A typical set of results from this exercise is shown in Table 1, generated by the World Bank’s PovcalNet, which in turn implements the method used by Chen and Ravallion, updating the information periodically.

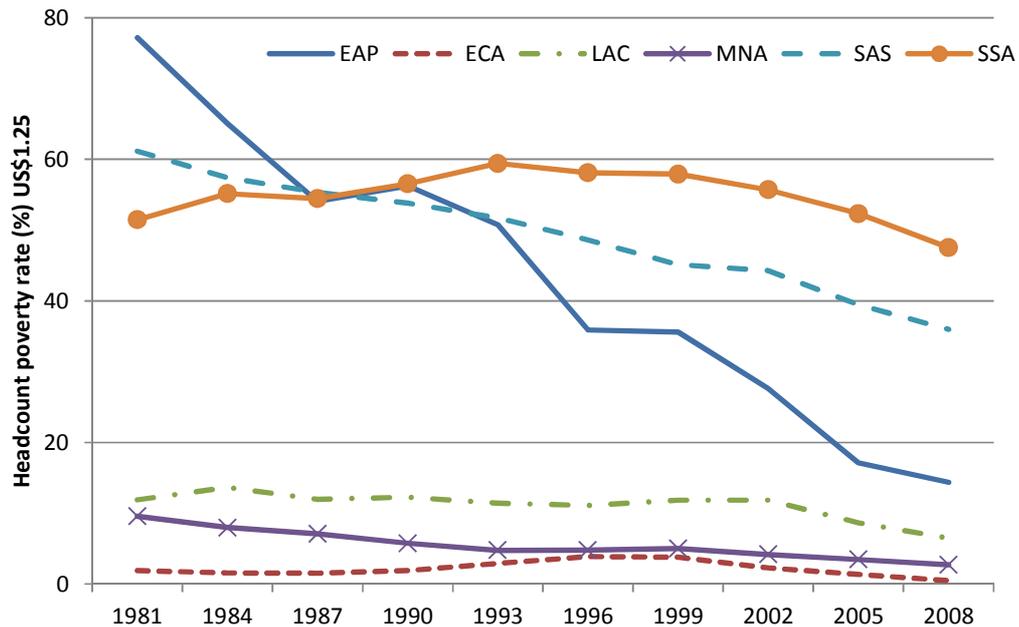
We see clearly from Table 1 that the percentage of people in less-developed countries living on less than US\$1.25 per person per day fell by more than half between 1981 and 2008, from 52% to 22%. The poverty rate fell dramatically in East Asia after 1981, mainly due to economic growth in China. Until the late 1990s, poverty rates continued to rise in Sub-Saharan Africa, but since then they have been falling there, and in every other major region of the world. Using a poverty line of US\$2.00 a day, poverty also fell over time, although less dramatically than the extreme poverty represented by the US\$1.25 per day standard.

The method used by Chen and Ravallion is the only practical approach to measuring worldwide poverty, but it is not without its critics. The PPP exchange rates should, in principle, reflect just the cost of living of the poor, not of the population in general, but this is not done in practice. The measures of PPP exchange rates are themselves subject to measurement error, and are not very stable over time (Pogge and Reddy 2003). Household survey data typically underestimate consumption and income, which will typically overstate the poverty rate. And the US\$1.25-a-day poverty line does not have a solid foundation in welfare theory; it does not, for instance, seem to track the cost of basic needs very well (Reddy et al. 2006). Despite these problems, the most satisfactory alternative – asking each country to measure poverty using a cost-of-basic-needs approach – is still impractical. So in what follows we use, *faute de mieux*, the estimates based on the Chen and Ravallion methodology.

**Table 1.** Percentage of Population in Less-Developed Countries Living on less than US\$1.25/person/day

	1981	1984	1987	1990	1993	1996	1999	2002	2005	2008
East Asia and Pacific	77.2	65.0	54.1	56.2	50.7	35.9	35.6	27.6	17.1	14.3
<i>of which:</i> China	84.0	69.4	54.0	60.2	53.7	36.4	35.6	28.4	16.3	13.1
Europe and Central Asia	1.9	1.6	1.5	1.9	2.9	3.9	3.8	2.3	1.3	0.5
Latin America & Caribbean	11.9	13.8	12.0	12.2	11.4	11.1	11.9	11.9	8.7	6.5
Middle East & N. Africa	9.6	8.0	7.1	5.8	4.8	4.8	5.0	4.2	3.5	2.7
South Asia	61.1	57.4	55.3	53.8	51.7	48.6	45.1	44.3	39.4	36.0
<i>of which:</i> India	59.8	55.5	53.6	51.3	49.4	46.6	44.8	43.9	41.6	34.5
Sub-Saharan Africa	51.5	55.2	54.4	56.5	59.4	58.1	57.9	55.7	52.3	47.5
<b>Less-Developed Countries (LDCs)</b>	<b>52.2</b>	<b>47.1</b>	<b>42.3</b>	<b>43.1</b>	<b>41.0</b>	<b>34.8</b>	<b>34.1</b>	<b>30.8</b>	<b>25.1</b>	<b>22.4</b>
<b>LDC population, million</b>	<b>3,715</b>	<b>3,943</b>	<b>4,185</b>	<b>4,433</b>	<b>4,665</b>	<b>4,891</b>	<b>5,115</b>	<b>5,328</b>	<b>5,537</b>	<b>5,746</b>
<i>Memo Items (all LDCs):</i>										
US\$1/day poverty line	41.6	34.7	30.1	30.8	28.6	23.5	23.1	20.6	16.0	14.0
US\$2/day poverty line	69.6	67.9	64.7	64.6	63.0	58.6	57.4	53.4	46.9	43.0
US\$2.50/day poverty line	74.9	74.2	72.2	71.7	70.6	67.5	66.2	62.5	56.6	52.4

Source: PovcalNet; accessed July 20, 2012.



**Figure 1.** Headcount Poverty Rates over Time for Six Regions  
*Notes.* EAP: East Asia and Pacific. ECA: Eastern Europe and Central Asia. LAC: Latin American and the Caribbean. MNA: Middle East and North Africa. SAS: South Asia. SSA: Sub-Saharan Africa.

**Picturing Poverty: Snapshots**

Table 1 is informative, but not entirely satisfactory. For instance, it does not tell us where most of the world’s poor live; nor does it help us understand the possible association of poverty with income levels. To do this, we could present more tables that summarize data; alternatively, in the spirit of Alan Gelman’s admonition to make more use of visual data exploration (Gelman et al. 2002), we could develop the analysis using graphs.

A good place to start is with a line graph that shows the poverty rate for each region over time (Figure 1). Although somewhat cluttered, the key messages jump out more easily here than they do from Table 1: rapidly falling poverty in East Asia, a moderate drop in South Asia, and little change in Sub-Saharan Africa until after 1999.

A two-way graph is more useful when we are trying to associate one variable with another, as in Figure 2, where the headcount poverty rate in 1981 (left panel) and 2008 (right panel) is shown on the vertical axis, and Gross Domestic Product (GDP) per capita, in PPP dollars on the horizontal axis. The link between poverty and income (as measured by GDP/capita) is striking: not surprisingly, as income rises, poverty falls. The most interesting anomaly here may be the relatively high poverty rate, given its income level, for Latin America and the Carib-

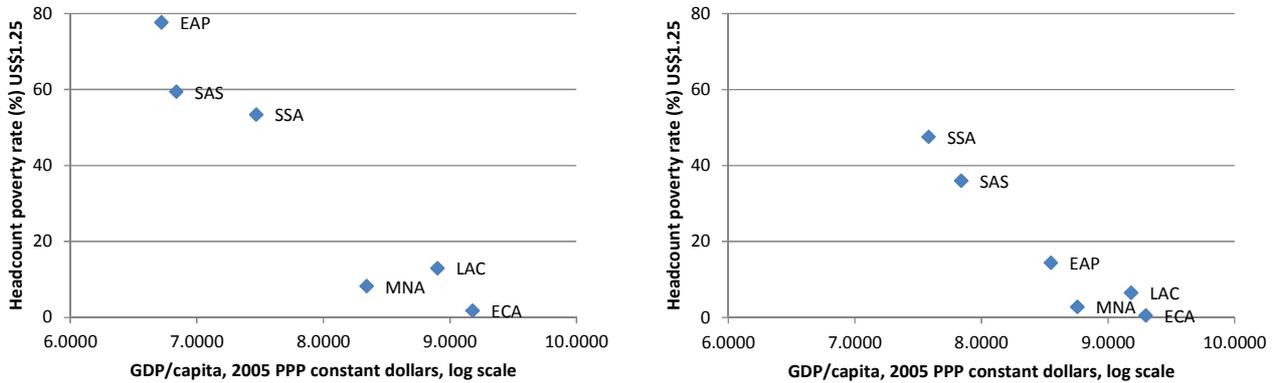
bean, which is a region where most countries have high levels of inequality internally.

Figure 3 represents an improvement on Figure 2, although the two graphs have the same variables on the X- and Y-axis. The difference is that each bubble in Figure 3 shows the *number* of poor people. This makes a big difference to our appreciation of the extent of poverty: In 1981, the greatest pool of poor people was in East Asia, mainly China (the large red bubble in the left panel), but by 2008 the number of poor East Asians had dropped sharply. On the other hand the number of poor Africans rose sharply over the period, the result of a rising poverty rate in the first two decades of the period under consideration, coupled with a rapidly-rising population.

**Picturing Poverty: Movement**

The bubble plots in Figure 3 could be made better yet if they were animated, so that we could see the changes in poverty rates, the number of poor, the location of poverty, and average incomes, over time. That is five dimensions at once!

Hans Rosling had developed some wonderful animated bubble charts, which he has used to communicate his ideas with real flair. On his Gapminder site (<http://www.gapminder.org/>) one can create moving



**Figure 2.** Associating Poverty Rates with GDP/capita for Major Regions, 1981 (left) and 2008 (right)  
 Source: Table 1 for poverty rates, and World Bank (World Economic Indicators) for GDP and population data.  
 Notes. EAP: East Asia and Pacific. ECA: Eastern Europe and Central Asia. LAC: Latin American and the Caribbean. MNA: Middle East and North Africa. SAS: South Asia. SSA: Sub-Saharan Africa.]

charts for the data series that are posted there, but those series do not include poverty rates.

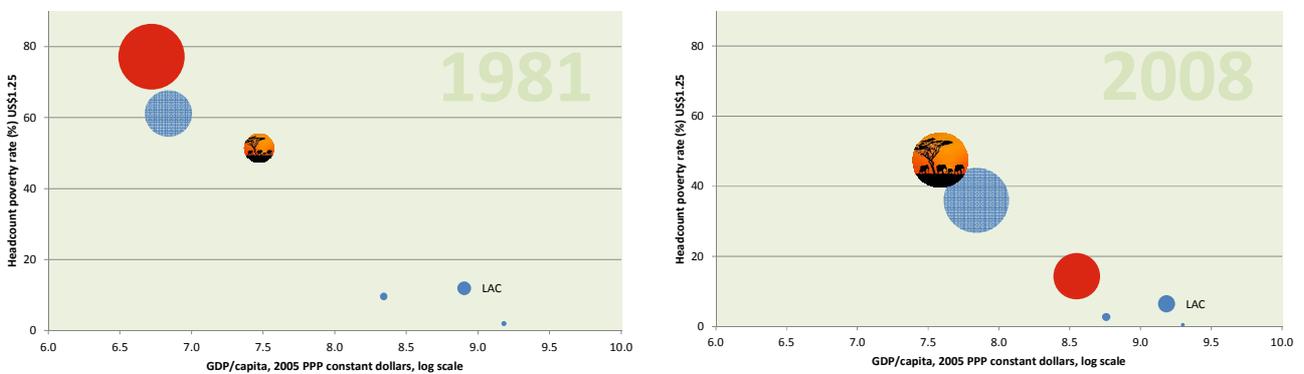
An animated bubble chart can be constructed fairly easily in Excel, following the steps outlined below. To see our bubble plot in action, you can find the file at <http://web.cas.suffolk.edu/faculty/jhaughton/BubblesGDPp125r2012.xlsm> (and it is also provided with this paper); then just click on the AnimatedGraphs tab, and you will get a spreadsheet with this graph; click on the button labeled “Play” and the graph will move, or use the slider at the right of the graph to move it one year at a time.

**The Mechanics of Constructing an Animated Bubble Plot in Excel**

The instructions here rely heavily on the excellent posting by S. Anand (<http://www.s-anand.net/blog/motion->

[charts-in-excel/](#) ). It is also possible to make dynamic bubble charts using the tools available in Google Analytics – for instance at <http://www.google.com/ig/directory?type=gadgets&url=www.youcalc.com/apps/1238149110412/igoogle.xml>. We use Excel because it is widely employed, so the learning curve for most people will be shorter.

First, organize the data as shown at the top of the next page, with a column for the date, the region (“Thing”), the log of GDP per capita (“X”), the poverty rate (“Y”), and the number of poor people (“Z”), in millions. In addition create a variable called “Key” that concatenates the date and region; for instance, the command could be “=A6&B6”. Note that I have created measures of the poverty rate for every year, by linearly interpolating for the years that are not included in the PovcalNet computations.



**Figure 3.** Associating Poverty Rates and Totals with GDP/capita for Major Regions, 1981 (left) and 2008 (right)  
 Source: As for Figure 2.

Date	Thing	Key	X	Y	Z
1981	EAP	1981EAP	6.72	77.18	1087.38
1981	ECA	1981ECA	9.18	1.91	6.08
1981	LAC	1981LAC	8.90	11.89	46.44
1981	MNA	1981MNA	8.34	9.56	13.86
1981	SAS	1981SAS	6.84	61.14	546.02
1981	SSA	1981SSA	7.47	51.45	213.40
1982	EAP	1982EAP	6.76	73.13	1030.59
1982	ECA	1982ECA	9.15	1.79	5.64

Note: EAP: East Asian and Pacific. ECA: Europe and Central Asia. LAC: Latin America and the Caribbean. MNA: Middle East and North Africa. SAS: South Asia. SSA: Sub-Saharan Africa.

Second, label a cell with a word such as “Offset”, and put in an initial value of 0, as shown here; be sure to name the cell by typing the name Offset over the band on the top left that gives the cell address. Then create a cell called “Today” and put in the initial data plus the value in offset, for instance with the command “=A6+Offset”. Cell A6 has the initial date.

	I	J	K	L
	Offset	0		
	Today	1981		
	Thing	X	Y	Z
6	EAP	6.7216	77.18	1087.4
7	ECA	9.1804	1.91	6.1
8	LAC	8.9034	11.89	46.4
9	MNA	8.3422	9.56	13.9
10	SAS	6.8392	61.14	546.0
11	SSA	7.4717	51.45	213.4

Third, create a lookup table like the one shown at the bottom of this page. In our example, the margins (Thing, EAP, X, etc.) are typed in directly. The numbers in the table are created by looking up the large dataset. The table itself will display the numbers, such as 6.7216 (EAP row, X, 1981), but the numbers will be drawn from the full dataset by using the “=vlookup(.” command, as shown in the figure below; it requires four arguments – the first is the value to look up (e.g. 1981EAP), the second is the location of the full data set, the third is an index of the column in the dataset from which one

	I	J	K
	Offset	0	
	Today	=A7+Offset	
	Thing	X	Y
6	EAP	=VLOOKUP(\$J\$3&\$I6,\$C\$7:\$F\$156,2,FALSE)	=VLOOKUP(\$J\$3&\$I6,\$C\$7:\$F\$156,3,FALSE)
7	ECA	=VLOOKUP(\$J\$3&\$I7,\$C\$7:\$F\$156,2,FALSE)	=VLOOKUP(\$J\$3&\$I7,\$C\$7:\$F\$156,3,FALSE)
8	LAC	=VLOOKUP(\$J\$3&\$I8,\$C\$7:\$F\$156,2,FALSE)	=VLOOKUP(\$J\$3&\$I8,\$C\$7:\$F\$156,3,FALSE)

should draw the value, and the last is a logical operator that is set equal to “FALSE” if the lookup match is to be exact.

The fourth step is to create a bubble chart based on the data in the lookup table. Highlight the numbers in the lookup table, and click on: Insert – Charts – Other Charts . Then select the plain version of the bubble charts. Make the graph look good! Here are a few thoughts on this process:

- The bubbles may be too big. Solution: Right click on any bubble, select “Series Options”, and scale the bubble size.
- We don’t know which bubble refers to which region. Solution: Color the bubbles differently. Click on a bubble until just that bubble is highlighted. Then right click, and choose “Format Data Point.” If you want to color the bubble, pick “Fill”, check “Solid fill”, and choose a color. It is also a good idea to set the transparency to about 30%, so that you can partly see through a bubble, which is helpful if there are other bubbles behind it. If you want to get fancy, you can put a picture on a bubble – as I have done for the Africa bubble. Instead of picking “Solid fill”, choose “Picture or texture fill”, and then insert the picture you want from a file or Clip Art. Be warned, however, that this can clutter a chart very easily, so don’t be tempted to over-decorate the graph, as that could detract from its central message.
- The bubbles may be lost in the graph. Solution: limit the axes. It is always a good idea to label the axes so that the graph is essentially self-contained.
- Adding a date in the background. Solution: Click on a bubble, right click, and click on “Select data.” In the pop-up window, click on Edit, and in the box that says “Series name:” put in the spreadsheet address of the year (for instance J3). Click on OK as required. Now click on the chart, click on the “Layout” tab under “Chart Tools” at the top of the screen, and click on “Chart Title” and choose “Centered Overlay Title.” The year will now appear on

chart. Format it to your taste: click on the title until it is shown in a box, then highlight the date, and make it big – I use a 72 point font – and then color it in a faint but visible hue. It will appear in the background, visible but not intrusive.

- Label the bubbles. One can also put labels next to the bubbles. Click on the chart, then pick Layout (under Chart Tools) and Data Labels. Determine where you would like the data labels to be. This will put numbers next to each bubble. Now go to the data label on each bubble, click twice, and type the label you want (e.g. SSA for Sub-Saharan Africa) in place of the number. Again, be careful not to clutter the graph unduly. I have included a label for Latin America and the Caribbean, by way of illustration; this is a small bubble, for which coloring does not work well.

Once you are satisfied that you have a beautiful bubble chart – my version, after considerable experimentation, is in Figure 3 – the fifth and final step is to animate it, which may be done in two ways, with a manual scroll bar, and with an automated button.

To add a scroll bar: if you have Excel 2007, click on the Developer tab, and then click on Insert, and click on the icon (under “Form Controls”) that gives a scroll bar. This will be a vertical bar with a slider, and you can format it to look nice next to the bubble plot. Now right click on the scroll bar, go to Format Control, click on the Control tab, and link the scroll bar to the Offset cell. Now as you move the slider on the scroll bar up or down, or click on its arrows, the graph will change from one year to the next.

To add a “Play” button: Using Excel 2007, click on the Developer tab, then Insert, choose the icon for “Button (Form Control)”, accept the prompt for a macro, and place the button where you want to put it. You can rename the button by right clicking on it, choosing “Edit Text”, and typing “Play” or whatever name you want to call it.

Now right click on the button, and choose “Assign Macro”, which will bring up a small window. Click on “New”, and when the appropriate box appears, type in the following:

```
Declare Sub Sleep Lib "kernel32" (ByVal
dwMilliseconds As Long)

Sub Button1_Click()
Dim i As Integer
For i = 0 To 27:
Range("J2").Value = i
Application.Calculate
```

```
Sleep (500)
Next
End Sub
```

You may need to modify this code somewhat. If your button has a different number you might need to put Button2\_Click, for instance. The number 27 reflects the number of years for which we have data – in this case, 1981 through 2008 – minus 1. The “J2” refers to the address of the “Offset” cell, which starts at 0 (for 1981) and rises to 27 (for 2008). And the number 500 in the “Sleep” line controls the speed with which the bubble plot will move; higher numbers will slow the animation down.

Now just return to the spreadsheet; when you click on the “Play” button, the graph will be animated. If you have other graphs that are linked to the same numbers, then will be animated too, as in our example.

### Lessons from the Animation

After going to the trouble of creating the animated bubble graph, what do we learn? The answer is that several useful themes emerge. First, the great bulk of world poverty is in East Asia, South Asia, and Sub-Saharan Africa; the number of poor people in the other parts of the world is relatively small. A second storyline is that Africa, which was less poor than South and East Asia in 1981, is now the poorest region; this helps explain both the renewed interest in poverty reduction in Africa, and the importance of the upsurge in African economic growth over the last several years. The third clear pattern from the animated picture is that East Asia has been transformed: half of the world’s poor lived there in 1981, but by 2005 it was clear that serious poverty in the region was rapidly disappearing. Finally, there is some evidence from the graph that since about 2000, all regions have been moving steadily towards the southeast – in other words, have experienced higher incomes and lower poverty rates.

The improvements that have occurred at the broad regional level have not been mirrored in every country. For instance, Haiti and Zimbabwe have become poorer than they were twenty years ago. It would therefore be valuable to disaggregate these graphs further; we leave this exercise to the enthusiastic reader.

Correspondence: [jhaughton@suffolk.edu](mailto:jhaughton@suffolk.edu)

## REFERENCES

- Balassa, B. 1964. "The Purchasing Power Parity Doctrine: A Reappraisal", *Journal of Political Economy* 72(6): 584–596.
- Chen, Shaohua, and Martin Ravallion. 2001. "How Did the World's Poorest Fare in the 1990s?" *Review of Income and Wealth*, 47(3): 283-300.
- Chen, Shaohua, and Martin Ravallion. 2004. "How Have the World's Poorest Fared Since the Early 1980s?" Development Economics Working Paper No. 3341, World Bank, Washington DC.
- Chen, Shaohua, and Martin Ravallion. 2008. "The Developing World Is Poorer Than We Thought, But No Less Successful in the Fight against Poverty," Policy Research Working Paper No. 4703, World Bank, Washington DC.
- Gelman, A., C. Pasarica, and R. Dodhia. 2002. "Let's Practice What We Preach: Turning Tables into Graphs," *The American Statistician*, 56(2): 121-130.
- Haughton, Jonathan, and Shahidur Khandker. 2009. *Handbook on Poverty and Inequality*, World Bank, Washington DC.
- Pogge, Thomas, and Sanjay Reddy. 2003. "Unknown: The Extent, Distribution, and Trend of Global Income Poverty." Columbia University, NY.
- PovcalNet: the on-line tool for poverty measurement developed by the Development Research Group of the World Bank.  
<http://iresearch.worldbank.org/PovcalNet/index.htm>  
[Accessed July 20, 2012]
- Reddy, Sanjay, Sujata Visaria, and Muhammad Asali. 2006. "Inter-Country Comparisons of Poverty Based on a Capability Approach: An Empirical Exercise." Working Paper No. 27, International Poverty Centre, UNDP, Brasilia.
- Samuelson, P. A. 1964. "Theoretical Notes on Trade Problems", *Review of Economics and Statistics* 46(2): 145–154.
- World Bank. 2000. *World Development Report 2000/2001: Attacking Poverty*. Washington DC.