Homework #1: Drivers of Change
(due Tuesday, Sept. 4th in class)

This assignment is intended to accomplish two things: (1) to give you a sense of the demographic and technological drivers of change over time and differences across countries in agricultural markets around the world; and (2) to give you practice in collecting and analyzing data from authoritative sources.

The first part of the homework aims to get you started with some data on world agriculture as a whole, asking questions of interest to the whole class. The second part asks you to find and analyze additional data of particular interest to you. These instructions offer detailed guidance on how to complete Part 1, so you can do that quickly and see how to make similar computations on your own for Part 2.

Obtaining original data and transforming it into meaningful charts or tables is a fundamental skill for professional economists – after this homework, I hope you will be more readily able to make your own analyses from original source data, instead of using only the charts and tables that others have made. Whenever you see a reference to data, you should be able to download the original variables and make your own customized charts and tables. This is a key step in transitioning from being a *consumer* to being a *producer* of economics research.

Part 1. Common data and analyses
The data for part 1 are all computed and published by the FAO in its FAOSTAT database, which is the most authoritative source of comparable aggregate data on food and agriculture across countries. These data are publically accessible at **http://faostat3.fao.org**. Everyone in the class will be downloading the exact same data, but you should be able to do this work independently on your own computer in preparation for Part 2 of this homework; to speed things I up I have provided some detailed step-by-step instructions below.

Before you start, note that you will be downloading the FAO's estimates, for each major region of the world, of four key numbers:

(i) an estimated number of farmers, which they call **employment in agriculture**;
(ii) an estimate of total farm output, which they measure using a value-weighted net production index (**Net PIN**);
(iii) an estimate for the area of useful farmland, which is **arable land**; and
(iv) an estimate of key inputs used, for which we count **nitrogen fertilizers**.

To see the FAO's definitions of these variables, from **http://faostat3.fao.org** you can click on **Methods & Standards** and then review the methodologies, classifications, units, glossary and abbreviations, as needed. To see the FAO's description of how they computed each variable, look on the **Methodology** to find a variable's name and click through for a description of how each estimate was obtained. Some of this is quite tedious, but it is good practice to know what your data *actually* represent!

To interpret our data, we will use them to compute four important ratios: available land per farm worker, total output produced per worker, total output produced per area of land, and total fertilizer use per area of land. Our goal is to compare these ratios across each of five major regions: Africa, South America, South Asia, Southeast Asia, and the world as a whole. Two of our ratios have natural units involving a number of hectares, workers or kilograms, and so we can compare the absolute level of these ratios as well as their year-to-year changes. To describe total output, we need an index number, for which only year-to-year changes are meaningful. Most data should be complete through 2016.

To obtain the data and make needed computations, step-by-step instructions follow. Note that from time to time the FAO takes down data for revision and some series may be unavailable. <u>If you are having trouble getting a full series, work with a partial series. If data are completely unavailable, check with me</u>.

(1) From **http://faostat3.fao.org** choose Browse (to look at the data) or Download (to build your own database for the assignment). Note that some items may be accessed differently depending on whether you are working in the Browse or Download mode. You may need to adapt these instructions at times. Click on "**Production**" in the drop-down menu and then select "**Production Indices**" from the second-level menu.

(a) The appropriate filter is **Regions**. Select the following regions (to highlight multiple entries, hold down the **"Ctrl"** key while clicking on each one):

World + (Total)
Africa + (Total)
South America + (Total)
Southern Asia + (Total)
South-Eastern Asia + (Total)

(b) Under **Items Aggegated,** choose **Agriculture (PIN) +** and under **Element,** choose **Net PIN (base 2004-2006=100),** then select all available years (to highlight a series of entries, hold down the **"Shift"** key while clicking at the start and end of the series):

(c) You can **preview** to see the results, but should choose **download** (and CSV under the drop-down tab) to build your spreadsheet. You will obtain a file with a name ending in **.csv**, which you should open with **Microsoft Excel** and then save to your hard disk, under a new name: I would suggest something like "**FAO-AgProdIndex.xls**".

(2) Now return to the top level menu and click on **Inputs**, then choose **Land** and repeat steps (a)-(c) above, but under **Items,** choose **Arable Land** and then save the result under a different name such as "**FAO-ArableLand.xls**".

(3) Still under **Inputs,** click on **Fertilizers** and again repeat steps (a)-(c) above, but under **Items,** choose **Nitrogen Fertilizers (N total nutrients)** and under **Element,** choose **Consumption in nutrients**, and save the result under a different name such as "**FAO-FertUse.xls**". Note that fertilizer data may be available for fewer years than production.

(4) Still under the **Inputs** tab, choose **Employment Indicators** followed by **Employment in agriculture**, and save the result under a different name such as "**FAO-EmpAgr.xls**".

(5) Now you can begin to use these data, by combining them into one file. You can move each spreadsheet into a common workbook, by right-clicking on the tab at the bottom of each sheet and choosing "**Move or Copy**" to append each sheet onto a single file that you might save under a name such as "**AGEC640-hw1.xls**", or you can use Windows to cut and paste all four blocks of data directly onto one sheet and then save that file under the new name.

(6) Like almost all real-world data, you need to think carefully about what the numbers mean, and transform them into some kind of ratio or index number to make economically meaningful comparisons. Then, to visualize and communicate the results, you need to present the data in carefully formatted line graphs that reveal trends and fluctuations over time. In this case, I would like you to create four charts whose titles are listed below:

Fig. 1. FAO estimate of arable land per agricultural worker (ha/persons), 1961-2016
Fig. 2. FAO index of net agricultural output per agricultural worker (PIN base 2004-2006), 1961-2016
Fig. 3. FAO index of net agricultural output per area of arable land (PIN base 2004-2006), 1961-2016
Fig. 4. FAO estimate of fertilizer use per area of arable land (kg/ha), 2002-2016

Figures 1 and 4 are ratios of numbers that have natural units (hectares, persons or kilograms), but figures 2 and 3 require the use of an index to add up quantities of

disparate things such as bulk grains, liquid milk, vegetables and flowers. We want to see how that index changes relative to the number of workers and the area of arable land. The ratio of an index to a number is itself an index, and to see how it changes over time please re-scale it to 2004=100. (Note that the actual index is scaled to a three-year average, 2004-2006=100, but you will obtain a similar result with just the year 2004=100.) Thus, the formulas you need to compute in Excel can be written as:

(1) $AgLandPerAgWorker_{year\_t} = AgLand_{year\_t} / AgWorkers_{year\_t}$

(2) $AgOutputPerAgWorker_{year\_t} = 100*(AgOutput_{year\_t} / AgWorkers_{year\_t})/(AgOutput_{2004}/AgWorkers_{2004})$

(3) $AgOutputPerArableLand_{year\_t} = 100*(AgOutput_{year\_t} / ArableLand_{year\_t})/(AgOutput_{2004}/ArableLand_{2004})$

(4) $FertUsePerArableLand_{year\_t} = FertUse_{year\_t} / ArableLand_{year\_t}$

You can do these computations quickly in Excel. For example, with your variables in rows and a column for each year, for charts (1) and (4) you can create and copy the same formula over all regions and years, while for (2) and (3), to make the index relative to 1961, insert a dollar sign ($) in front of the column letter denoting 2004 for each variable in the denominator before you copy that formula to all other years. Note that mistakes in Excel formulas are very common, so check your work very carefully at each step.

To turn your data into a nice chart, an easy approach is to put all variables in a block with years across the top and regions on the left, and then left-click in one corner and drag your mouse across the entire block; once all data are highlighted, you can build your chart. Stata makes particularly attractive figures, so if you know Stata then you might want to copy and paste from Excel and work there. Regardless of what you use to generate your figures, choose a **line** chart showing straight lines (not smoothed) between data points, and make sure that any missing data are represented as blank cells (rather than zeros).

Format your lines and markers denoting each region so that your chart can be easily read when printed in black and white. Label each axis, and make other edits as needed to obtain a clear, clean figure. The goal should be to produce the cleanest and most easily read figure possible. This is not an assignment for an Art class, so avoid using fancy fonts, shadows, outlines, etc.

(7) Once you have made nice charts in Excel or Stata, please cut-and-paste them into a Word document.

(8) Finally, please interpret what these data tell you about the evolution of agriculture in each region and in the world as a whole. I expect you to write several paragraphs under each graph to describe what you see and how that might relate to the data on your other graphs. Note that a key point from the Week 2 readings is that the trends in Figure 1 (land area per worker) are largely driven by factors

outside the farm sector, driven by the demographic transition and structural transformation of the economy as a whole. They are the demographic and structural “drivers of change”, to which the agricultural sector then responds. From that figure you should be able to explain in which regions and at what times farmers have been driven to look for ways of using land more intensively (that is, applying more labor per acre), and when/where they have been driven to look for ways of using land more extensively (that is, applying their labor to more land).

(9) When discussing Figures 2, 3 and 4, please explain the degree to which farmers in each region might have succeeded in raising output per worker, output per acre and input use per acre. What might drive the trends you see, particularly the relative speed and timing of changes in these trends? What kinds of policies might have helped make some regions more successful than others? Be careful about assigning direct causality, but feel free to speculate.

Part 2. Individual data and analyses

The data and analyses described above are drawn from FAOSTAT, which is the authoritative source on global agriculture; next you should turn to the most authoritative source of more general data across countries, which is the World Bank's "World Development Indicators" (WDI). Unlike FAOSTAT this requires a subscription, so should be accessed through the Purdue library by choosing it from the alphabetical list in the drop-down **Find** menu at the top of **http://www.lib.purdue.edu/mel/.**

Your task in this part of the homework is similar to Part 1, except that you are now free to choose whichever countries (or regions) and whichever variables are of most interest to you.

Please produce exactly four additional charts analyzing some important trends, and then explain those charts as you did the figures described above. Note that the WDI offers access to a very wide range of data, although most variables turn out to be available only for certain countries or certain years. The data you choose do not have to be available for all years, but please ensure that missing observations are shown as blanks rather than zero.

You may wish to look at different data for similar regional aggregations as in Part 1, or you may wish to look at data for individual countries and specific aspects of government policy of greatest interest to you. Note that the World Bank's definition of each region may be different from the FAO definition. Click on the "info" or "help" buttons for guidance.

In writing up your analysis of the four original charts in Part 2 of the assignment, please be sure to explain (a) why you chose those data, (b) what you think are plausible relationships between them and (c) how they relate to the process of agricultural development illustrated in Part 1 of this assignment.

I would be happy to discuss the assignment as you proceed. If you have a technical question about the workings of FAOSTAT or the World Development Indicators, it is almost certainly best for you to call me (494-4218) while you are at the computer working on the database, so that you can describe exactly what you see.