
Empirical Assignments

We are going to be working with a wide range of evidence in class related to different dimensions of economic development. Whether looking at contract terms written on cuneiform tablets, transcripts of criminal trials at the Old Bailey, or modern national accounting data, every data source offers a unique perspective on the evolution of the standard of living and has its own unique advantages and disadvantages. These assignments will give you the chance to grapple with a wide range of data sources and critically assess just how much they can and cannot tell us about differences in economic development across time and space. Beyond helping you develop a more nuanced understanding of the data underlying the papers we cover in class, these exercises will also provide practical experience with tracking down a variety of data sources and presenting them in an effective manner.

Each assignment will typically require you to track down a specific type of evidence, create an effective visualization of those data, and offer a paragraph of interpretation. The timing of each assignment will be roughly matched to the schedule of lecture content, giving you a chance to take a deeper dive into the concepts we discuss in class. Each assignment will push you in different directions in terms of the types of data and types of visualizations you will need to tackle.

The Assignments

1. Measuring the standard of living

At its heart, this class is about exploring how we shifted from a world of stagnant (and low) standards of living to a world with steady increases in the standard of living for some countries but continued stagnation for others. In this assignment we will dig deeper into just how we measure the standard of living and what we miss with those measures. The goal of this assignment is to assess how modern nations compare along two different dimensions of well being. More specifically, you need to do the following:

1. Find two different measures of the standard of living for modern countries. One measure should directly relate to economic well being (income, wealth, consumption expenditures, etc.) while the other should relate to physical well being (mortality, general health, etc.). You need to put together a single cross section of data: one observation of your economic well being variable and one observation of your physical well being variable per country. Not all observations need to come from the same year (both within and across countries) but all should be modern (no earlier than 2010).
2. Construct a scatterplot with the measure of economic well being on the horizontal axis and the measure of physical well being on the vertical axis. Each point on this plot will

correspond to a country. You may not be able to get data for all countries but your figure should have points for at least 100 countries. Be certain to make your graph as clear as possible: use an informative title, appropriate axis titles, and figure notes as appropriate.

3. Write one paragraph that addresses the following questions:
 - (a) To what extent do the two measures seem well correlated?
 - (b) What aspects of the standard of living does your economic well being measure likely miss?
 - (c) What aspects of the standard of living does your physical well being measure likely miss?
4. Submit your figure, paragraph, and full citations to your data sources as a single pdf document through Blackboard.

2. Geographic endowments and economic development

One of the theories we cover in class is that of [Diamond \(1999\)](#), outlined in his book *Guns, Germs and Steel*. A key part of his argument is that geographic endowments matter. If you have natural endowments of useful plants and animals, or if you live in an area where useful plants and animals can spread to, you get a leg up when it comes to economic development. One aspect of this theory is that being on a useful latitude matters while being on a particular longitude does not: climates are similar along latitudes meaning that plants and animals can thrive at similar latitudes but might not thrive when moving north and south along a particular longitude. In this exercise, we will see if basic economic data seem to support this theory. Here is how you will do this:

1. Create a dataset that includes the latitude and longitude of every country (as with the first assignment, you may not be able to find data for all countries but should be able to get data for at least 100 countries). Obviously, countries are not single points so there will be different ways to define a single latitude and longitude for each country, just be certain to apply a consistent method across all countries. Note that you should not have to track down the latitude and longitude of each country individually, you should be able to find a data source that gives you latitudes and longitudes for all countries.
2. Now find data on the economic development of each country. The variable you choose is up to you but it must: (i) be from the 21st century, (ii) be the same variable for all of your countries, (iii) clearly be a reasonable proxy for economic development and (iv) be different than the variables you used in the first assignment. Merge these data with your latitude and longitude data for each country.

3. Create two scatterplots, one showing the relationship between your proxy for economic development and the latitude of a country and another showing the relationship between your proxy for economic development and the longitude of a country. Make the figures as effective as possible by making appropriate choices regarding titles, labels, axes and so on.
4. Write one paragraph discussing whether your figures support or contradict Diamond's theory.
5. Submit your figures, paragraph and full citations to your data sources as a single pdf document through Blackboard.

3. Property rights and investment

Several of the theories of industrialization we reviewed in class emphasize the importance of growth-promoting institutions generally and secure property rights specifically (North & Thomas, 1970; Acemoglu et al., 2002, 2001). For modern countries, economists use a variety of ways of quantifying the strength of property rights. This can include looking at the laws related to contracts, piecing together measures of corruption, interviewing business leaders as well as a number of other approaches. This exercise will explore whether good institutions are indeed correlated with economic development in a modern context by having you complete the following:

1. Track down a country-level measure of growth-promoting institutions for the continent of your choice (excluding Antarctica and North America). This could be something related to control of corruption, strength of property rights, or a range of other things. You should strive to get data for all countries on the continent, though one or two may not have data available. Note that if you choose Australia, you should get data for the other island nations in Oceania as well, not just Australia. As with the other assignments, the expectation is that you will be able to find a single data source with data on multiple countries; you should not need to investigate each country individually.
2. Merge these data with data on foreign direct investment in each of the countries in your dataset. Increased foreign direct investment is one of the clearest indicators that businesses feel secure in their investments in a country. You only need one year of foreign direct investment for each country. While the year you use can vary across countries, try to be as consistent as possible. You would not want to compare foreign direct investment during the Great Recession for one country to that for another country during an economic boom.
3. Create two maps, one showing the variation in growth-promoting institutions across your continent and a second depicting the variation in foreign direct investment. Note that you may want to rescale or otherwise normalize your data to make the maps clearer. Make certain the maps have clear legends.

4. Write one paragraph interpreting your maps, assessing whether they do or do not reveal a relationship between institutions and investment. Write a second paragraph that focuses on the effectiveness of a map versus a scatter plot for presenting the information. This second paragraph should be informed by both your maps for this assignment and the scatterplots you made for the previous assignment. Be certain to consider the advantages and disadvantages of both types of figures.
5. Submit your maps, paragraphs and full citations to your data sources as a single pdf document through Blackboard.

4. Time use and the changing work week

One remarkable feature of the Industrial Revolution was the change in work hours of individuals. In class, we covered the clever work of [Voth \(2000\)](#) using court records from the Old Bailey to piece together the work schedules of witnesses. In modern studies, we have the advantage of surveys of both firms and workers providing detailed accounts of hours worked. In this assignment, we will explore the use of firm-level original records to fill the gap between Voth's estimates of hours worked during the Industrial Revolution and our modern estimates from the Bureau of Labor Statistics and others. This will require some work with primary sources, specifically the ledgers and payroll accounts of farms, mills and other employers in Virginia throughout the 1800s and early 1900s. These data sources give us an amazing chance to uncover far more detailed information on work and wages than the census occupation data relied on by the vast majority of American economic history studies. Here is the plan:

1. We will have a class visit to the Special Collections on Thursday, March 28th. This will take place during our usual class period. During this visit, the librarians at the Special Collections will present you with a range of historical documents from the collections. You will get to work with six different ledgers that span a wide range of years and types of employers. The goal is to use these ledgers to explore a rich variety of information about the nature of the work week.
2. Your group (which we will assign in advance) will start at one of the ledgers. Take your time to turn through the pages, getting a sense of how the ledger is set up, how the information varies from page to page and the range of information it contains. Once you have perused the ledger, choose one page that you think contains a good range of information about occupations, wages and/or work hours. Photograph the page, ensuring that your digital images are good enough to be able to read all of the text on the page. Also take an image of the slip indicating the library catalog details of that particular ledger so you know which images correspond to which ledger.
3. Every few minutes, we will rotate groups so that your group will get a chance to spend time with each of the six ledgers. For each ledger, do the same as what you did with the first. Take a few minutes to look through the ledger and then choose a page with

significant information on occupations, wages, and/or work hours and take high-quality photographs of the page.

4. At the conclusion of the Special Collections visit, each group will have images from six pages, one from each ledger. As a group, you will digitize the information from three of these six pages. It is your choice as to which three pages to digitize. When digitizing the pages, you should create a separate spreadsheet for each one. You should structure your spreadsheet in a way that captures the structure of the ledger page and then record all of the information on the page. If any characters are truly illegible, record them with an asterisk (*). For example, if my name was listed as one of the employees, and you could make out my first name but only first three letters of my last name, you would record it as "John Parm***". Note that you can split the digitization work up however your group decides. You can work on all three pages as a group, have each group member digitize half a page, or any other approach that works best for you. For each of your spreadsheets, make certain that the title of the file indicates which ledger the page is from and, if it was given on the page, the page number.
5. Individually (not as a group), write one paragraph that discusses one aspect of the information in the ledgers that surprised you (this can be from any of the six ledgers you looked at, not just the three your group digitized). Explain what was surprising about it and why it might be of interest to economic historians (think about what sort of paper idea your discovery might motivate). Write a second paragraph that discusses how using historical ledgers like those in the Special Collections might give us biased estimates of either work hours or wages. This second paragraph does not need to be motivated by anything specific you noticed in the ledgers, though you can certainly draw on specifics from the ledgers.
6. On Blackboard, submit your six ledger images, your three spreadsheets of digitized ledger data, and a pdf containing your two paragraphs of discussion. Note that the six images and three spreadsheets will be identical for everyone in the same group but each individual should still submit them with their assignment. The paragraphs of discussion will be different for everyone in the group. For the images and spreadsheets, any file formats are fine, though the spreadsheets should be in an editable format (Excel, Google Sheet, csv, etc.).

5. Fear and faith in technological change

More than anything else, this class is about technological change and its impact on the typical worker. While we have largely focused on sustained technological change driving sustained growth in the standard of living, workers often fear being displaced by technology. From the machine breaking by Luddites during the Industrial Revolution to the handwringing over advances in artificial intelligence in recent years, technological change has routinely altered the landscape of work in complex ways and garnered strong responses both positive

and negative. While we will dig into the economic theory of this in class, exploring work like [Goldin & Katz \(1998\)](#) on skill-biased technological change, we are going to take a very different approach in this exercise. The main goal is to see if sentiment analysis of texts can tell us about how the general public thought of technological change over the past two centuries (recall [McCloskey \(2006\)](#) on attitudes toward economic activity in literature, a related approach). Here is how you will do this:

1. Find three different passages from novels, opinion pieces, magazine articles or any other non-academic source that discuss a major technological change. One of your passages should be written in the 19th century, one in the 20th century and the final one in the 21st century. Each passage should be one paragraph long.
2. For each passage, assess how positive or negative you find the overall tone of the passage. You should assess this on a scale from zero to ten, with zero being completely negative and ten being completely positive. The idea here is to try to quantify overall sentiments toward technological change.
3. Now go to the [Google Form](#) for submitting responses and enter in all of the relevant information. You will need to enter the paragraphs themselves, your score for the overall sentiment and citations for each paragraph. When the form is completed, you should receive a copy of your responses via email. If you do not, let me know and we can figure out what went wrong.
4. Once I have received everyone's responses, I will aggregate the passages into a single database and run it through a sentiment analysis program to estimate an additional dataset of sentiment scores. We will then spend time in lecture comparing how your sentiment scores compare to those of the algorithm and use that as the foundation of a discussion around the possibilities and limitations of automated text analysis to analyze history.

Assignment Due Dates

Below is a table giving due dates for each assignment. Assignments are due by **5pm** on the date listed. Assignments turned in after the due date will incur a small deduction in the grade of one point (assignments are graded out of 20 points). That deduction will increase by one point after every additional 48 hours. The goal of this grading policy is to strike a balance between incentivizing you to meet deadlines, which in turn keeps your workload for this class balanced across the semester, and making certain that unexpected obstacles that lead to late assignments do not severely impact your grade.

Assignment	Due date
Measuring the standard of living	February 8
Geographic endowments and economic development	February 22
Property rights and investment	March 21
Time use and the changing work week	April 11
Fear and faith in technological change	April 25

Potentially Useful Links

- [Google Books](https://books.google.com): <https://books.google.com>

Google Books provides a wealth of digitized, searchable texts ranging from novels to government reports. If you click on ‘Search tools’ you will be able to narrow your search to specific dates of publication, to only documents fully available online, and to specific types of documents. Particularly if you are looking at older time periods, Google Books may be a valuable source for the text of political speeches or for published polling data.

- [Google Scholar](https://scholar.google.com): <https://scholar.google.com>

Google Scholar has become the easiest way to search for scholarly articles on any subject. Simply search for academic research in the same way that you would search for anything else on Google. When you find a relevant article, Google Scholar provides several nice features. It can often provide a link to a pdf version of the article, it provides a link to the article in the Swem library database, it can show you articles citing the current article and it provides simple tools for copying citation details for the article.

- [NBER US Business Cycle Expansions and Contractions](http://www.nber.org/cycles.html): <http://www.nber.org/cycles.html>

This website lists the peaks, troughs and durations of US business cycles back to 1854. When identifying periods of strong and weak economic growth for your country of choice, you will want to find a similar source of dates. Note that Wikipedia can be a useful starting place when searching for this type of information. For example, these NBER US business cycle dates are referenced on the Wikipedia page [List of recessions in the United States](#).

- [World Bank Open Data](http://data.worldbank.org/): <http://data.worldbank.org/>

This is the data portal for the World Bank. You can find a broad range of variables for different countries through this site. Students often find the [World Development Indicators](#) particularly useful. These data include detailed statistics on the population, economy, environment, markets and global links of individual countries.

- Various government statistics websites

Some of the best sources of data for a country are the governmental agencies responsible for gathering and publishing economic statistics. In the United States, these agencies would include the [Census Bureau](#), the [Bureau of Labor Statistics](#), the [Bureau of Economic Analysis](#) and many others. If you go to any of these websites, you find links to detailed datasets. Searching for comparable government agencies for your country of focus will likely lead to a wealth of official statistics.

- [Purdue Online Writing Lab](https://owl.english.purdue.edu/owl/section/2/): <https://owl.english.purdue.edu/owl/section/2/>

This online writing lab provides helpful guides to citation styles including APA style, MLA style and the Chicago Manual of Style. Note that the site also provides useful guides for conducting research and general writing.

- [William & Mary Libraries](https://libraries.wm.edu/): <https://libraries.wm.edu/>

Always remember that we have an excellent library with excellent librarians ready to help you. Research librarians have extensive knowledge about available data sources and can help point you down the right path, making the research process far more efficient. You can even click [here](#) to make a research appointment.

- [Wikipedia](https://www.wikipedia.org/): <https://www.wikipedia.org/>

While Wikipedia should never be your final source for information, it is often an excellent initial source for information, particularly for dry economics topics that do not attract contentious edits. A quick trip to Wikipedia can often help you get your bearings, identify key terms to search for on Google Scholar, and provide a few initial citations to track down as a good start to your research.

- [Datawrapper](https://www.datawrapper.de/): <https://www.datawrapper.de/>

While the majority of students will make their figures in Excel, a perfectly fine option, a relatively new online option for constructing very effective graphs is Datawrapper. Access to the resource is free and the default formatting of figures is far nicer than Excel, while still giving you plenty of options to customize figures. Datawrapper also has a very intuitive interface for making maps, something that anyone who has ever struggled with GIS software in the past will appreciate.

- Articles, books and blogs on representing data

There are many articles, books and blogs out there focused on the visualization of data and in particular the visualization of economic data. For a true classic, take a look at [Tufte \(1983\)](#). More recent takes on data visualization can be found in [Chen et al. \(2007\)](#) and [Wong \(2010\)](#) while a short guide more specific to the economist is provided by [Schwabish \(2014\)](#). Note that Schwabish also has a book specifically on data visualization in Excel which is available [online](#) from the William & Mary library ([Schwabish, 2023](#)). You may also find inspiration from the variety of data visualization blogs out there including [Flowing Data](#) and [Junk Charts](#).

References

- Acemoglu, D., Johnson, S., & Robinson, J. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *The American Economic Review*, 91(5), 1369–1401.
- Acemoglu, D., Johnson, S., & Robinson, J. A. (2002). Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *The Quarterly Journal of Economics*, 117(4), 1231–1294.
- Chen, C.-h., Härdle, W. K., & Unwin, A. (2007). *Handbook of data visualization*. Springer Science & Business Media.
- Diamond, J. M. (1999). *Guns, Germs, and Steel: The Fates of Human Societies*. Norton New York.
- Goldin, C., & Katz, L. F. (1998). The Origins of Technology-skill Complementarity. *The Quarterly Journal of Economics*, 113(3), 693–732.
- McCloskey, D. (2006). The Discreet virtues of the Bourgeoisie. *History Today*.
- North, D. C., & Thomas, R. P. (1970). An Economic Theory of the Growth of the Western World. *The Economic History Review*, 23(1), 1–17.
- Schwabish, J. (2023). *Data Visualization in Excel: A Guide for Beginners, Intermediates, and Wonks*. CRC Press.
- Schwabish, J. A. (2014). An economist’s guide to visualizing data. *Journal of Economic Perspectives*, 28(1), 209–34.
- Tufte, E. (1983). *The visual display of quantitative information, Second Edition*. Graphics Pr.
- Voth, H.-J. (2000). *Time and Work in England 1750-1830*. Clarendon Press Oxford, UK.
- Wong, D. M. (2010). *The Wall Street Journal guide to information graphics: The dos and don’ts of presenting data, facts, and figures*. WW Norton.