1. Goals of Assignment

In Assignment 2, we will apply the econometric tools learned throughout the course to <u>extend the analysis</u> of a published paper in economics. While Assignment 1 acquainted you with the main building blocks of coding and analyzing data to answer a specific research question, Assignment 2 will expand on that process offering a little more flexibility in working with data and little more flexibility in defining the analysis. This is an incremental approach to preparing you to write subsequent papers based on your own ideas and data collection in other courses (e.g. Eco 475). Particularly, if you intend to go on to write empirical projects in the future, you might take this "replicate and extend" strategy as a good first pass way to find a feasible research idea in your topic of interest. The overarching goal, of course, is to give you the building blocks to constructing formal analysis in your future careers. [1]

Similar to Assignment 1, your final analysis will be structured and organized like a short paper. The template is the same here, and you are still provided a context and data upon which to base your submission. Furthermore, the assignment requirements and scope are still somewhat tightly defined (see below) meaning there is less leg work in refining your own question and less range for missing the mark in your execution.

The data and focus of the Assignment 2 this year is the relationship between income and democracy using data from Acemoglu, et al. (2008). A complete description of the data and topic is given in the Context and Data section below. This is a case where subsequent team of authors took the "replicate and extend" approach and were able to publish, years later, a comment on the original paper in the *American Economic Review*. So, as you can see, such a thing is possible.

2. The Context and Data

Do citizens demand more democracy and political freedom as their incomes grow? This is the question we will be grappling with in Assignment 2 using data from Acemoglu, et al. (2008). A quick scan of the relationship comparing countries (i.e., a simple cross-sectional regression across countries), yields a positive correlation between income per capita and political freedom. But what underlies this relationship? Is the correlation explained by a causal connection whereby a higher income per capita causes a country to become more democratic, or is it explained by the reverse: democracy, itself, causes higher incomes per capita? Lastly, nothing rules out the idea that some third factor is a determinant of both higher incomes and higher political freedom; a so-called omitted variable (which we know all about by now).

The core focus of your assignment will be to analyze the relationship between income and democracy using the panel dataset from Acemoglu, et al. (2008). Of course, more details on this topic can be found through a literature search (which will obviously include Acemoglu, et al. (2008) in the least). Your job is to study this question and analyze this relationship empirically, and your goal is to evaluate this paper and extend this analysis further in some regard.

Data: I provide a random sample of 147 countries from the five-year panel in Acemoglu, et al. (2008)

---

[1] The "replicate and extend" strategy can be a helpful way to cut through <u>immediately</u> to feasible ideas in your field of choice (which is especially useful if your timeframe is a semester or two in length). This is because when you consider the full set of interesting research ideas, you'll find the feasible ones are usually (and unfortunately) a small subset. And when you are working within a tight timeframe, you don't want to waste your efforts cycling through too many "pie in the sky" notions. The replicate and extend strategy can be useful here as a way to narrow focus, but also as an information gathering exercise; it helps you become aware of data choices, estimation strategies, and interpretation issues in the specific field, which can eventually build nicely into your own independent analysis. Eventually, you may just get to that pie in the sky, but you'll do so by meeting all your deadlines along the way.

on the course website, and you will keep to this set of 147 countries for all of your analysis (base specification and extensions).[2] This means that if you add to or extend the dataset (which is not necessarily required), please make sure you stick to the same 147 countries (for consistency with the base specification results). As well, if you decide to explore the 10-year panel (or beyond) from Acemoglu, et al. (2008), you need to merge these data into the Assignment 2 data and keep only the matched countries for analysis. The variables in the dataset are described below and also described in more detail in Acemoglu, et al. (2008).

Variable	Description (see also Acemoglu, et al. (2008))
code	Country code
country	Country name
age_midage	Percent population age 30-45
age_old	Percent population age 45-60
age_veryold	Percent population age 60-
age_veryyoung	Percent population age 0-15
age_young	Percent population age 15-30
code_numeric	Country numerical code
education	Average schooling years
fhpolrigaug	Augmented Freedom House Political Rights Index
laborshare	% labor share of gross value added
lpop	Log(total population in thousands)
lrgdpch	Log real GDP per capita (PWT)
medage	Median age in population
nsave	Nominal savings rate: (Y-C-G)/Y
polity4	Polity IV index
socialist	Dummy for Soviet Block, including iron curtain
year	Year of observation
year_numeric	Year of observation numerical
sample	Dummy for base sample
worldincome	Trade weighted world income instrument
worlddemocracy	Trade weighted world democracy
samplebalancefe	Dummy for balanced sample for fixed effects
samplebalancegmm	Dummy for balanced sample for GMM
colony	Dummy former colony vs non-colony (from: CEPII http://www.cepii.fr)

3. Structure of Assignment

The finished assignment will be comprised of a concise abstract; tables: (1) descriptive statistics, and (2) regression results; and no more than 6-7 pages (1.5-spaced) of corresponding text (in 12-point font) outlining and interpreting the empirical results. This should be organized as follows:

Page 1: Title page
Including course and student information, and a concise abstract.
Page 2-7 (or 8): Text
Formal academic composition with a 6-7-page limit, 12-point font, and 1.5 line spacing. The text should NOT include tables or figures.
Page 8 (or 9): References
Includes a list of all references cited in your assignment.

---

[2] This means, even if you run an identical regression, your estimates will not be numerically equivalent to the tables in Acemoglu, et al. (2008), and your course knowledge from the early week's lectures tells you why: because our estimators are random variables. Nevertheless, we can argue that our estimates are estimating the same parameters as in Acemoglu, et al. (2008) because the countries in our "daAssignment2.dta" sample are missing at random, meaning we can argue that the countries that are missing should not be systematically related to our variable(s) of interest.

Page 9 (or 10) - : Tables and figures

Tables and figures are appended after the list of references. All included tables and figures must be discussed in the text. DO NOT include undigested STATA output! Results presented in tabular form should have all variables and numbers clearly labeled. Any figures should also be well labeled and clear. Consider each Table/Figure as a stand-alone product, which can be largely understood on its own (i.e. without referring to the written text).

Course Tools:

Overall, the assignment should demonstrate application of tools discussed throughout the course, but especially those from week 7 onward. Background reading includes Acemoglu, et al. (2008) and potentially follow up literature. You should notice already, a difference in your ability to read such articles and recognize the underlying econometric structure as applied to the context. In the very least, we will achieve this at the end of the assignment. This is our last hurrah. In service of this, you will want to consider the various types of specification types and issues we have covered since chapter 5, and since we are dealing with a panel dataset, Chapter 13 and 14 are also useful background resources. Lastly, Chapter 19, section 19-5 "Writing an Empirical Paper" in the course textbook has some general ideas about style guidelines in economics writing.

4. Content requirements

Following the overall structure outlined in Section 3, your assignment should read like a short paper. The content and scope in this assignment should follow the guidelines below.

4.1 Overall Scope:

4.1.1   Base specifications: your assignment will include analysis and write up on the regression of the democracy index (i.e., *fhpolrigaug,* sourced from the Augmented Freedom House Political Rights Index) on the five-year lag of log real GDP per capita (i.e., the lag of *lrgdpch,* sourced from the Penn World Tables). As stated above, you will run your analysis on the dataset provided on the website called "daAssignment2.dta", which is a random sample of 147 countries from the five-year panel dataset from Acemoglu, et al. (2008). Since the base specifications are tightly defined, you can view this part of the assignment like a tutorial problem where you are given the data and asked for a specific type of analysis. Details on the base specifications will be given in section 4.2 below.

4.1.2   Extension: your assignment should include an extension to the base analysis in 4.1.1, where the choice of extension is up to you. Acemoglu, et al. (2008) describes the data and provides several appendix tables detailing the sources (Appendix Tables A1-A2). Consider our course tools and possible ways to use them to extend the analysis; the choice is up to you. Some examples might include exploring non-linear effects in income, exploring differences in the effect with respect to a third variable, extending the time period to see if results maintain, etc.

It is fine to estimate specifications suggested by the subsequent literature, but you must keep to the overall sub-sample of 147 countries we are using in the Assignment 2 dataset. Of course, nothing precludes you from running analysis on subsets of this sample of countries if this is your direction of choice. The reason we have limited the analysis to this random sample of countries, is so the estimation is comparable to the results in the base specification. This means that, if you use 10-year panel (or beyond), make sure your results are on the sample of 147 countries as the base specification sample. If you update the dataset, then once again, make sure the estimation sample is limited to the same sample of 147 countries as the base specification. Lastly, if you do end up adding to the data, you must document this addition, appendage or merge in your code file, and upload the modified/additional data with your submission so we can replicate your results after submission.

4.2 Detailed Guidelines

4.2.1   Title Page: this will include course and student information, a title, and a concise abstract

4.2.2   Abstract: a brief summary of your findings (on the title page)

4.2.3   Introduction: introduce and motivate the analysis.

4.2.4   Describe the data: describe the data, referring to, at least, a table of descriptive statistics.

- This table will be named Table 1 and will be included after the references page.
- It will contain pertinent summary statistics like the means, s.d.'s, frequencies, etc. for the variables used in the analysis.
- Means and standard deviations allow you to put the size of estimated coefficients in context. A well-written analysis will discuss summary statistics as a prelude to the main point of the analysis and will use summary statistics to put the size of estimated effects in context (e.g. what is the average of the democracy index? how do summary statistics vary across high income and low income countries, across colonies vs non-colonies? etc.). These examples are just ideas, of course, the structure and expositional choices are up to you.

4.2.5   Regression analysis section:

Base Specifications: your assignment will include initial analysis and write up of regressions of the democracy index sourced from the Augmented Freedom House Political Rights Index on the five-year lag of log real GDP per capita: i.e., *fhpolrigaug* on the one period lag of *lrgdpch*. Last time, I created all lag variables for you, but this time this will be up to you. To create the lag of any variable in Stata you can use the lag operator: e.g., gen lag_varx=L1.varx, where "varx" is the original variable, "L1." is the lag operator for one period, and "lag_varx" is the new variable. You can also put lag operators directly into your regression as in: "reg fhpolrigaug L1.lrgdpch".

Analyze the following specifications and report them in Table 2 (included at the end of your assignment). A template for Table 2 is shown after this description, and you should report all statistics shown in the template even if not explicitly mentioned below.

1. A SLR of *fhpolrigaug* on the one period lag of *lrgdpch* with robust standard errors.
   - Interpret the estimated coefficient on the lag of *lrgdpch*? Is it large or small? Is it statistically significant? Report and interpret the 95% confidence interval for this estimate.
2. A SLR of *fhpolrigaug* on the one period lag of *lrgdpch* with standard errors clustered at the country level.
   - Does the estimate change?
   - Report and interpret the 95% confidence interval for this estimate. Why is it important to use clustered standard errors for the regression?
   - Does this suffer from omitted variable bias? Provide an example of a time invariant factor that varies with democracy and income systematically, and may bias results.
3. A regression of *fhpolrigaug* on the one period lag of *lrgdpch* including country fixed effects.
   - Interpret the estimated coefficient on the lag of *lrgdpch*? Is it large or small? Is it changed from (2)? Is it statistically significant? Report and interpret the 95% confidence interval for this estimate.
   - Does this still suffer from omitted variable bias? We are controlling for fixed country specific factors here, but what about time varying confounders? Can you provide an example of a time varying factor that varies little by country, which may also vary with democracy and income systematically?
4. A regression of *fhpolrigaug* on the one period lag of *lrgdpch* including country and year fixed effects.
   - Interpret the estimated coefficient on the lag of *lrgdpch*? Is it large or small? Is it changed from (3)? Is it statistically significant? Report and interpret the 95% CI for this estimate.
   - Test the exclusion of the year fixed effects and report the F-statistic and p-value.
5. First run a regression of *fhpolrigaug* on the one period lag of *lrgdpch* including country and year fixed effects, and lagged demographic controls: age, education, and log population (you will have to choose which age category to leave out). There is missing data in the demographic controls, so we are going to use this regression to define a subsample. After any regression, Stata records a postestimation sample indicator in e(sample). Create a variable say: "gen demo_sample=(e(sample)==1)", then run specification (4) on this subsample. Report results.

6. Now run a regression of *fhpolrigaug* on the one period lag of *lrgdpch* including country and year fixed effects, and lagged demographic controls: age, education, log population (you will have to choose which age category to leave out).
   - Interpret the estimated coefficient on the lag of *lrgdpch*? Is it large or small? Is it changed from (4)? Is this because of the change of sample? Is it statistically significant?
   - Test the exclusion of the year fixed effects and report the F-statistic and p-value.
   - Test the exclusion of the age variables and report the F-statistic and p-value.
   - Test the exclusion of all the demographic controls and report the F-statistic and p-value.
   - Interpret the estimated coefficients on the demographic controls and the F-statistics.

Table 2: Regression Analysis…

	(1)	(2)	(3)	(4)	(5)	6
Lag of log real GDP pc	estimate	estimate	estimate	estimate	estimate	estimate
	(s.e.)	(s.e.)	(s.e.)	(s.e.)	(s.e.)	(s.e.)
	95% CI	95% CI	95% CI	95% CI	95% CI	95% CI
Age						
Age group: …						estimate
						(s.e.)
Age group: …						estimate
						(s.e.)
Age group: …						estimate
						(s.e.)
Age group: …						estimate
						(s.e.)
Education						estimate
						(s.e.)
Log Population						estimate
						(s.e.)
F-statistics and P-values						
Year Fixed Effects				estimate	estimate	estimate
				(p-value)	(p-value)	(p-value)
Age Controls						estimate
						(p-value)
Demographic Controls						estimate
						(p-value)
Year Fixed Effects	No	No	No	Yes	Yes	Yes
Country Fixed Effects	No	No	Yes	Yes	Yes	Yes
Clustered Standard Err.	No	Yes	Yes	Yes	Yes	Yes
Demographic Sample	No	No	No	No	Yes	Yes
Number of Obs.	value	value	value	value	value	value
Number of Countries	value	value	value	value	value	value
R-Squared	value	value	value	value	value	value

Notes: pertinent details; source data; details on variable definitions; time period; etc.

***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

Extension: your assignment will extend this analysis by exploring several different specifications of your choice, which you will report according to your own expositional stylings.

Discussion of results: be sure to interpret key coefficients, as well as the results of the hypothesis tests. For example, in your discussion of results be sure to mention:

- The size of the estimates in practical terms.

- When adding a set of control variables, discuss whether the estimated coefficient change in terms of both: statistical significance and practical significance.

- Discuss what happens when moving to from specification to specification.

4.2.6 Limitations of results: discuss possible problems with these specifications, especially omitted variables that may still be an issue (i.e. do you interpret your results as causal or are they purely descriptive?). What are the most important remaining threats to the validity of your regression results?

4.2.7 Conclusion: Based on your analysis what conclusions would you draw about the effects of income on democracy.

5. Evaluation

There is no mechanical grading scheme. But the following should be a helpful guide and explain in more detail what we are looking for in your submission.

5.1. Exposition and Writing (Approximately half the weight)

You will need to explain your empirical results to your reader. This involves describing the data; motivating the regression; clearly explaining what is estimated; posing and interpreting results. You may wish to consult Chapter 19, section 19-5 "Writing an Empirical Paper" in the course textbook for general ideas about style guidelines in economics writing. Key elements of clear exposition include:

1. Placing results in context (i.e. discuss "economic" significance)
   - Discuss up front (in introduction) and come back to it later (in the results section)
2. Data description (concerning Table 1)
   - Does anything hint at the subsequent empirical results?
3. Base Specification: explain, interpret
   - Definition of variables
   - Specify and motivate hypotheses
   - Map coefficients into interpretation
   - Discussion of key regression assumptions and how they are addressed
4. Extension
   - Motivate additional regressions and hypotheses
   - Reflects back on the base specification
5. Quality of the tables
   - We will penalize (heavily!) undigested STATA output (i.e. tables do not look like STATA output).
   - Tables should (ideally) have "English" variable names, not acronyms
   - Ask yourself: "If I had never seen this table before would I still understand what it is about?"
6. Integration of empirical results from tables to text.
7. Overall writing and coherence
   - Does the assignment draw results together?
   - Is it clear to read?

- Usually, a confusing assignment reflects underlying confusion.

5.2. Content (Approximately half the weight)

It is not easy to separate exposition from content, but in general characteristics we will be looking for execution of the content guidelines described in section 4.

1. Clear, succinct introduction.
2. Base specification: explanation, interpretation, discussion of results and assumptions
2. Estimation and discussion of additional regressions/extensions
   - Implementation of hypothesis tests and interpretation
   - When using dummy variables, make sure that the interpretation and specification is correct (i.e., relative to a base group)
3. Potential statistical problems
   - Be sure to discuss why/how a data problem might affect results (as opposed to just stating a problem might exist).

5.3. Due Date and Submission Details

Assignment submission has two main parts: Part 1: Crowdmark (details below) and Part 2: Quercus (details below). You will want to plan enough time to finish the submission process by the deadline as both parts of the submission are necessary components of a complete submission.

Why do we require so many files? Because it is good practice to document your work so that it might be replicated exactly and easily. This is not an ECO 375 thing; this is an economist thing. The journal websites we've seen from the *American Economic Review* and the *American Economic Journal: Applied Economics* provide good examples of how we document our estimations so that they might be estimated by others. You must look to do the same for your submitted work following the guidelines below.

5.3.1. Verify your Submission

You must verify (the correct version of) your work was submitted before the deadline. We will not grade the submission unless both submission parts are fully complete. If both submission parts are not completed in full by the deadline, a 100% late penalty will apply. Details on verifying your submission are in the links below.
- https://crowdmark.com/help/verifying-that-an-assignment-was-submitted/
- https://ctl1.utsc.utoronto.ca/quercus/help/Uploading_an_assignment_to_Quercus.pdf

If you anticipate difficulties, try using Chrome, clear your cache, restart your browser or computer, and close all other tabs and programs before submitting your work.

5.3.2. Part 1: Crowdmark submission

Location: Crowdmark: https://app.crowdmark.com/sign-in.
Due date: April 21st 10:00 am EST
Late Penalties: 100% penalty per 1-minute late.
File Format and Description:
- [required] Complete the written part of your assignment in Microsoft Word in .docx format. Convert the final version of this docx file into a pdf file and submit this pdf file to Crowdmark.

General Instructions:
- Follow all the instructions described in the Assignment Overview (this document).
- Submit to Crowdmark in the two prompts provided: Q1 will ask you to upload a scan of your Title page (page 1 as described above) with your student card overlaid at the top righthand corner. Q2 will ask you to upload your pdf file (all pages) to the system. Both of Q1 and Q2 are necessary to a complete submission. We will not grade your Quercus submission if the Crowdmark submission is not complete.
- Your answers must be TYPED. Any handwritten answers will get zero points.

- Verify your submission: https://crowdmark.com/help/verifying-that-an-assignment-was-submitted/ . This is especially relevant if you are having connection issues.

5.3.3. Part 2: Quercus submission

Location: Quercus Assignments: https://q.utoronto.ca/courses/197735/assignments/581935
Due date: April 21st 10:00 am EST
Late Penalties: 100% penalty per 1-minute late.
File Format and Description:
- [required] Submit the written part of your assignment in Microsoft Word docx format. This file must be the SAME <u>final version</u> docx file which produces the pdf submitted in Part 1.
- [required] Submit a Stata do file which contains the code that reproduces the results in your word file.
- [required] Submit a Stata log file that is produced by the submitted code file.
- [required in certain cases] If you have added to or modified the data you use in the extension section of results, please upload the dta dataset(s) so that we might replicate your results after submission. If you don't submit it, we will request it, and it will result in a grade of zero if results are not replicable.

General Instructions:
- Follow all the instructions described in the Assignment Overview (this document).
- This submission is to check/run your Stata support files and to run your work through Turnitin. This is also a necessary submission. We will not grade your submission in Crowdmark if this Quercus submission is not complete.
- If you ran your analysis in R, you can submit code and log files that reproduce your results in R. There is a wrapper to create a log file in R, which you may need to use.
- Your answers must be TYPED. Any handwritten answers will get zero points.
- Your answers must be in English without any other languages embedded in the file. Again, this will be an automatic zero if we cannot process the file through the required checks because of this.
- Verify your submission, including whether you have submitted the correct file: https://ctl1.utsc.utoronto.ca/quercus/help/Uploading_an_assignment_to_Quercus.pdf (scroll down to part 3). This is especially relevant if you are having connection issues.

5.3.4. Plagiarism

Plagiarism is a serious problem (in general) with university writing. Obviously, if we detect this form of academic dishonesty, we deal with it severely. Even "inadvertent" plagiarism is penalized. You should familiarize yourself with the rules regarding the citation of sources, etc.

5.3.5. Circumstances that will lead to a zero on the assignment

How to avoid a zero?
So, you want to avoid a zero on the assignment… here is a handy guide to doing so. First off, to avoid a zero, I suggest you read the other sections very carefully and adhere to all the requirements. This is what it takes. Secondly, to help you doubly avoid such a wretched thing, I want to reiterate a number of circumstances below that would lead to a zero on your assignment (this is not an exhaustive set of circumstances; there are, of course, many other creative ways to earn a zero).

Before I outline the circumstances, I want to point out that most students safely and easily avoid all scenarios that lead to zero on the assignment. I'm writing this section not because there is a high level of difficulty here requiring that I write a manual about it. Instead, I am writing this because some people try to avoid aspects of submission for various (sometimes nefarious) reasons, and I want to be very clear about the requirements to aid in any ex-post discussion about such things. Secondly, I also want to restate these requirements to reassure the <u>significant majority</u> of students who put effort into the assignment and work hard to meet the requirements by the deadline, that there is no backdoor way for others to get out of that work. Lastly, the policy adheres in <u>general terms</u>, but if you want an

example of the kind of cases these policies protect against, it could be something like a student who submits a plagiarized document to Crowdmark and tries to avoid detection by dodging submission of matching support files or a student who puts in an ex-post claim that they submitted the wrong version but waits until after solutions have already been released.

Circumstances where zero stands:

Incomplete Submission: incomplete submissions will be assigned a grade of zero. This means you must submit all parts of the assignment as outlined above, which includes the submission of four files as described in 5.3.2 to 5.3.3 and a fifth file under the conditions described in 5.3.3. "Files as described" means files are submitted that adhere to the rules in Section 5.3 above. For instance, a file will not count as submitted if your code file does not reproduce the results in your docx file, or your docx file is not what you submitted as a pdf, etc. We also do not count unrelated files. You can submit 4, 5, 6, 100 files; we will only count the files produced as described above for complete submission.

These files as described must be submitted by the deadline, and if all required files are not submitted by the deadline, you are at risk of a grade of zero. This is even where up to 3 of the files are submitted correctly, and it is definitely the case if either one of the docx or pdf files is missing, or they do not match each other. In these cases, your submission is not complete, we will NOT grade your submission, and the 100% late penalty will stand. To ward off this potential event, it is extremely important that you verify your files are submitted in the <u>correct format</u>, <u>correct version</u>, and <u>as required</u> above. This is YOUR responsibility. Note that we CANNOT submit your files for you as we DO NOT have access to your accounts and passwords. Note also, we DO NOT know which version you meant to submit, nor do we have access to your hard drive to search for it. This is why YOU need to verify YOUR submission.

Submission Errors: In some cases, a submission error occurs, and students promptly (and worriedly) inform us that this has happened. They send us <u>emails immediately</u> with <u>files attached</u> within <u>minutes</u> of the deadline. They know there is a problem because they experienced the submission difficulty for themselves, or they went to verify their files were submitted (as required) and discovered the error then. This kind of thing can happen, and it may not result in a 100% late penalty. This is because in the cases where students reach out <u>immediately</u> (i.e., within minutes) AND where we can <u>validate: the files, the timing, and the error</u>, we will override the late penalty. We are reasonable people. I'm not really talking about those cases here. I'm talking about the cases where the submission is not completed by the deadline and there is NO immediate follow up on the part of the student. And I am especially talking about the cases where there is NO follow up AT ALL until after GRADE and/or SOLUTION RELEASE.

To be very clear, an incomplete submission (see section above) means that you have NOT submitted the required files as described in section 5.3 by the deadline and you did not try to <u>immediately</u> address this issue when the deadline passed. Note also, we WILL NOT accept ignorance of incomplete submission as a defense. As belaboured above, verifying the files are submitted as described is YOUR responsibility. In this case, the 100% late penalty stands. Any ex-post inquiry into the 100% late penalty should expect a reply with a screen shot of this section of the overview with little explanation.

Heavy Reference of Other People's Work: If your writing heavily relies on quotations from other sources or heavily references other sources, we will penalize this, you guessed it: <u>heavily</u>. This is not your work, and we will not award you credit for it. For instance, if you quote or closely paraphrase the work of others in portions of your work (e.g., spanning numerous sentences), we will award you a mark of zero. This is even when you cite the source of the text (…of course, if you do not cite the source, this is plagiarism, and we will put it forth as an academic infraction).

On the other hand, if you quote or paraphrase others as a way to anchor the <u>addition</u> of your own analysis, this is fine. In this case, <u>most</u> of your writing will be about how <u>your</u> analysis reflects back on the original source analysis. In fact, this is not just fine, this is great! This is putting the "re" in "re" search.

References:

Acemoglu, Daron, Simon Johnson, James A. Robinson, and Pierre Yared. 2008. "Income and Democracy." *American Economic Review*, 98 (3): 808-42.DOI: 10.1257/aer.98.3.808