Useful Links

This document lists some useful links to find datasets.

Collections/repositories of datasets from multiple sources

Statistical database summary by the Marshal Library:
http://www.marshall.econ.cam.ac.uk/onlineresources/statistics
Contains links to a number of primary data sources such data on international trade, GDP, and results from national surveys. Most data is at the macro level, while it also contains links to restricted-access micro data (e.g. responses to surveys) for which you need your Raven id.

Princeton University guide to data sources
http://dss.princeton.edu/cgi-bin/dataresources/guides.cgi
Excellent web page with links to data on most thinkable subjects for your dissertation. Examples include micro data on consumer expenditures, daily inflation indices, surveys on medical expenditure, indices on all sorts of subjects like economic freedom or country risks, micro data on how people spend their time, etc. Note: this page is managed by the library of Princeton University and some links to datasets are only available to Princeton University students. If you need one of those datasets, you may find a copy through other sources, and the Princeton webpage can be a good place to start looking.

ICPSR (Interuniversity Consortium for Political and Social Research)
https://www.icpsr.umich.edu/icpsrweb/
This collection of data contains most of the US micro-datasets. It can be useful for studies and research in sociology, political science, economics, demography, education, child care, health care, crime, minority populations, aging, terrorism, substance abuse, mental health, public policy, and international relations.

WRDS (Wharton Research Data Services)
https://wrds-web.wharton.upenn.edu/wrds/index.cfm?
WRDS provides the user with one location to access over 250 terabytes of data across multiple disciplines including Accounting, Banking, Economics, Finance, ESG, and Statistics. WRDS provides access to S&P Capital IQ, CRSP, NYSE, Thomson Reuters, Bureau van Dijk, Global Insight, OptionMetrics and other important business research databases. Using a standard query structure, for example: company names, dates of interest, financial parameters, students can create customised reports. These reports can be stored, revisited and exported. The Information & Library Services team at Cambridge Judge Business School can authorise class accounts, to provide access for students on MPhil courses. We also have a class account that Economics undergrads are welcome to use. Before they first use it, students need to register to use Judge BS Library services at https://jbs.qualtrics.com/SE/?SID=SV_cT8ij1zSwBIQjFr
The library can then issue them with the username and password.

1 Updated on October 20, 2017. Suggestions are welcome: contact Maarten de Ridder, mcd58@cam.ac.uk or Alexei Onatski ao319@cam.ac.uk
Minnesota Population Center
https://www.ipums.org/
Contains micro datasets for the leading population surveys in the United States. The most important ones are the Current Population Survey and the American Community Survey. These contain survey responses from 40,000 and 300,000 (resp.) Americans on questions relating to income, schooling, education, family decomposition, unemployment, health status and insurance, etc. Data is easy to collect using menus and starts as early as 1850.
Note: you will need to make a (free) account with your Cambridge University e-mail and agree to terms and conditions (e.g. you will not use the data for commercial purposes) before you can download data. It may take a day for accounts to be approved.

UK Data Service Discover
https://discover.ukdataservice.ac.uk/
UK’s largest collection of UK and international social, economic and population data. Includes major UK government-sponsored surveys, cross-national surveys, longitudinal studies, UK census data, international aggregate, business data, and qualitative data. Students have to register, at no financial cost for access.

Macroeconomic data

European Central Bank’s Statistical Data Warehouse
http://sdw.ecb.europa.eu/
Macroeconomic and financial data for the Euro area.

St. Louis Fed
https://fred.stlouisfed.org/
Lots of freely downloadable macro and financial data

Federal Reserve Bank of New York:
Regional economic data; markets operation data; national economic data. Note also electronic data bases attached to published papers.

World Bank Open Data
http://data.worldbank.org/
Tons of data that can be browsed by country or indicator.

Bank of England Statistical Interactive Data base:
http://www.bankofengland.co.uk/boeapps/iaadb/NewInterMed.asp?Travel=NlxlRx
Economic, financial, interest rate and exchange rate data by: country, financial instrument, business category, economic and industrial sector. Note also electronic data bases attached to published papers.

Office for National Statistics (ONS)
https://www.ons.gov.uk/
UK’s largest independent producer of official statistics and its recognised national statistical institute. Lots of statistics related to the economy, population and society at national, regional and local levels.

United Nations Statistics Division: Statistical Databases
http://unstats.un.org/unsd/databases.htm
A variety of statistical resources of the UN system.

Bureau of Economic Analysis, US
http://www.bea.gov/
US National Accounts, International accounts (Aggregated), Regional and Sector Accounts, Input-Output data

US Census Economic Statistics:
http://www.census.gov/econ/
Detailed 5-year economics census data by industry, location, firm-size, firm-age, firm-location, small business data and summary statistics from annual surveys on firms.

http://www.bls.gov/data/
Detailed data on Employment, Unemployment, Inflation, Productivity, Wages, Occupational Statistics

EU KLEMS and World KLEMS (K-capital, L-labor, E-energy, M-materials, and S-purchased services
http://www.euklems.net/
http://www.worldklems.net/data.htm
Detailed Historical Statistics on Industry Accounts for large number of countries.

WIOD: World Input Output Data
http://www.wiod.org/new_site/data.htm
Series of input output tables giving trade across sectors both within and across countries as well as employment, capital, value added and environmental indicators at the industry-country level.

EUROSTAT
http://ec.europa.eu/eurostat/data/database
A comprehensive cross country coverage of macro, regional and industry accounts.

OECD
https://data.oecd.org/
OECD data

IMF data
The IMF publishes a range of time series data on IMF lending, exchange rates and other economic and financial indicators. The useful datasets available from this website include
International Financial Statistics (IFS), World Economic Outlook (WEO), Balance of Payments Statistics (BOPS), and Direction of Trade Statistics (DOTS).

Penn World Table
http://www.rug.nl/research/ggdc/data/pwt/
A database with information on relative levels of income, output, inputs and productivity, with country and period coverage depending on the release.

Microeconomic data

NATIONAL LONGITUDINAL SURVEY OF YOUTH 1979 AND 1997
http://www.bls.gov/nls/
Designed to document the transition from school to work and into adulthood. Collects extensive information about youths' labor market behavior and educational experiences over time. Also included is a survey of the biological children of women in the NLSY79. The 1979 survey began with over 12,000 participants, while the 1997 survey began with approx. 9,000.

PANEL STUDY OF INCOME DYNAMICS (PSID) SERIES. 1968+
http://psidonline.isr.umich.edu/
Designed to fill the need for a better understanding of the determinants of family income and its changes. Longitudinal survey of US individuals and the families in which they reside. Can be used for cross-sectional, longitudinal, and intergenerational analysis and for studying both individuals and families.

SURVEY OF INCOME AND PROGRAM PARTICIPATION. (SIPP) 1984+
http://www.census.gov/sipp/
Longitudinal U. S. government survey of the financial status of American households conducted since 1983 (data starts with 1984). Covers government transfer and service programs, pension coverage, housing affordability, home ownership data, housing cost data (primarily mortgages), financial assistance for education, among other topics. Data may also be accessed via ICPSR [http://www.icpsr.umich.edu/icpsrweb/ICPSR/series/135]. The National Bureau of Economic Research (NBER) [http://www.nber.org/data/sipp.html] makes the SIPP data and documentation available along with Stata programs for reading the data. Considered the most comprehensive survey for measuring household wealth by race and ethnicity.

BRITISH HOUSEHOLD PANEL SURVEY (1990-2009)
http://www.esds.ac.uk/findingData/bhpsTitles.asp
Main objective is to further understanding of social and economic change at the individual and household level in Britain (the UK from Wave 11 onwards), to identify, model and forecast such changes, their causes and consequences in relation to a range of socio-economic variables. Provides information on household organization, employment, accommodation, tenancy, income and wealth, housing, health, socio-economic values, residential mobility, marital and relationship history, social support, and individual and household demographics.
CONSUMER EXPENDITURE SURVEY (CEX, 1980-now)
http://www.bls.gov/cex/pumd.htm
The Consumer Expenditure Survey program consists of two surveys, the Quarterly Interview Survey and the Diary Survey, that provide information on the buying habits of America’s consumers, including data on their expenditures, income, and consumer unit (families and single consumers) characteristics. The survey data are collected for the Bureau of Labor Statistics by the U.S. Census Bureau. The CEX is the only Federal survey to provide information on the complete range of consumers’ expenditures and incomes, as well as the characteristics of those consumers. Importantly, it is the main dataset that includes information about both consumption and wealth, which is interesting for those that want to study the effects of inequality.

World Bank Micro Data Library
http://microdata.worldbank.org/index.php/home/
The Microdata Library facilitates access to data collected through sample surveys of households, business establishments or other facilities. These 'microdata' sets may also originate from population, housing or agricultural censuses or through an administrative data collection processes. The Library contains supporting documentation from censuses and surveys conducted or supported by the World Bank, as well as by other international organizations, statistical agencies and other agencies in low and middle-income countries.

Pew Research Center
http://www.pewresearch.org/data/download-datasets/
Survey data on American politics (requires signup but free otherwise)

Center for the Study of Auctions, Procurements and Competition Policy
http://capcp.psu.edu/data-and-software
Auction, procurement, and competition policy related data and software.

Micro data from developing countries
https://ipl.econ.duke.edu/dthomas/dev_data/
Extensive list of microdata sets on household and firm surveys for developing countries.

Add Health
http://www.cpc.unc.edu/projects/addhealth
Add Health combines US longitudinal survey data on respondents’ social, economic, psychological and physical well-being with contextual data on the family, neighborhood, community, school, friendships, peer groups, and romantic relationships, providing unique opportunities to study how social environments and behaviors in adolescence are linked to health and achievement outcomes in young adulthood. Some data are restricted-use. Public-Use data doesn’t contain ID numbers of friends, siblings or romantic partners, so the data cannot be linked.

Demographic and Health Surveys
Important financial databases

SNL Banking Data via Capital IQ
www.snl.com
SNL Financial provides data on approximately 29000 banking institutions worldwide, including ownership information, financial statements, market filings, merges and acquisitions’ information and financial sector news. US data going back up to 30 years, European data going back up to 13 years.

BLOOMBERG and DATASTREAM
(no online access)
Bloomberg provides current and historical financial quotes, business newswires, and descriptive information, research and statistics on over 52,000 companies worldwide. You only access the database at the terminal in the Marshall Library.

WRDS (Wharton Research Data Services)
https://wrds-web.wharton.upenn.edu/wrds/index.cfm?
WRDS provides the user with one location to access over 250 terabytes of data across multiple disciplines including Accounting, Banking, Economics, Finance, ESG, and Statistics. WRDS provides access to S&P Capital IQ, CRSP, NYSE, Thomson Reuters, Bureau van Dijk, Global Insight, OptionMetrics and other important business research databases. Using a standard query structure, for example: company names, dates of interest, financial parameters, students can create customised reports. These reports can be stored, revisited and exported. The Information & Library Services team at Cambridge Judge Business School can authorise class accounts, to provide access for students on MPhil courses. We also have a class account that Economics undergrads are welcome to use. Before they first use it, students need to register to use Judge BS Library services at https://jbs.qualtrics.com/SE/?SID=SV_cT8ij1zSw8IQiFr
The library can then issue them with the username and password.

ST LOUIS FED
https://fred.stlouisfed.org/
Lots of freely downloadable macro and financial data

FAMA-FRENCH Data library
http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
This site has a large amount of freely downloadable historical data compiled by Eugene Fama and Kenneth French. The data is updated regularly, and the Fama-French 3-factor data is especially useful for analyzing fund and portfolio performance.

Robert Shiller Library
This site has Robert Shiller’s PE10 data which was used in the book “Irrational Exuberance”. PE10 data is updated regularly.

**YAHOO FINANCE and GOOGLE FINANCE**
https://uk.finance.yahoo.com/
https://www.google.co.uk/finance?tab=ee
Provides historical price information at the daily frequency for thousands of assets, easy to download, no login etc

**ASSET MACRO**
http://www.assetmacro.com/market-data/
Historical Macroeconomic Data and Market Data (Stocks Indices, Bonds, Commodities, FX)

**OANDA**
A great resource for historical FX rates that you can easily download

**THINKNUM**
https://www.thinknum.com/
US macro data at the zip code level. Requires signup.

**GITHUB**
https://github.com/caesar0301/awesome-public-datasets#finance
This is an open source community for programmers. You can find lots of useful software and tools there, and also public datasets. Need to create an account, but free. https://vincentarelbundock.github.io/Rdatasets/datasets.html

**LOBSTER**
Trade and Quote data from NASDAQ
Available through Cambridge INET (contact Peter Malec pm563@cam.ac.uk)

**Caltech Quantitative Finance Group**
http://quants.caltech.edu/research.html
A wide selection of stock market data available for free

**Some commercial providers include**

**CME**
http://www.cmegroup.com/market-data.html

**Quandl**
https://www.quandl.com/
Includes some free data but need to signup

**Financial Times:**
http://markets.ft.com/data
Equities, currencies, commodities, bonds, world markets. Individual or corporate subscription for full access.

**Data used in various economic papers.**

Journal of Applied Econometrics Data Archive
http://qed.econ.queensu.ca/jae/
Contains data for all papers accepted for publication by the journal after January, 1994, unless the data are confidential.

Replication WIKI
http://replication.uni-goettingen.de/wiki/index.php/Main_Page
Links to thousands of Economic studies to replicate with references to data sources.

John Cochrane’s data
https://faculty.chicagobooth.edu/john.cochrane/
John Cochrane posts the data he uses in his papers, when appropriate, on his web page. He also provides his code. Many other economists do the same. It may be worth looking at the web pages of the economists who work on problems similar to yours.