Automating update of the Christiano, Motto and Rostagno (2014) database for the United States

Published on June 2, 2016 by Thomas Brand (/author/thomas-brand.html), Anatole Cheysson (/author/anatole-cheysson.html)

database (/tag/database.html) , model (/tag/model.html) , estimation (/tag/estimation.html) , R (/tag/r.html)

DOWNLOAD (HTTPS://GIT.NOMICS.WORLD/MACRO/CMR14-DATA)

Our purpose is to write a program to automatically update a database similar to the one used in the bayesian estimation of the DSGE model developed in Christiano et al. (2014) (hereafter CMR) for the United States.

We need 12 series for the estimation of the CMR model:

1. GDP
2. Deflator of GDP
3. Consumption
4. Investment
5. Deflator of investment
6. Wages
7. Hours worked
8. Loans to non financial corporations
9. Short-term interest rate
10. Net worth
11. Credit spread
12. Term premium

To those 12 series we added 2 others:

1. Loans to households and non-profit institutions serving households
2. House prices

To create this database, many sources are used (BEA, BIS, BLS, OECD). Mostly, we take data directly from DBnomics (https://db.nomics.world/), through the rdbnomics (https://cran.r-project.org/web/packages/rdbnomics/index.html) package. For Wilshire 5000 Total Market Index and Moody's Seasoned Baa Corporate Bond Yield, we use data from FRED (https://fred.stlouisfed.org/) (from Saint-Louis Fed) using the plugin function FredR (http://www.jankocizel.com/FredR/) from Janko Cizel, which is updated (https://github.com/thomasbrand/fredR). All the following code is written in R, thanks to the RCoreTeam (2016) and the RStudioTeam (2016).

Raw data from BEA, BIS, BLS and OECD

We use DBnomics to retrieve data from BEA, BIS, BLS and OECD.

```
df <- rdb(ids=c("BEA/NIPA-T10106/A191RX-Q",
                "BEA/NIPA-T10109/A191RD-Q",
                "BEA/NIPA-T10106/A006RX-Q",
                "BEA/NIPA-T10109/A006RD-Q",
                "BIS/total_credit/Q.US.N.A.M.XDC.A",
                "BIS/total_credit/Q.US.H.A.M.XDC.A",
                "BIS/selected_pp/Q.US.N.628",
                "BLS/pr/PRS85006033",
                "BLS/pr/PRS85006103",
                "OECD/MEI/USA.IRLTLT01.ST.Q",
                "OECD/MEI/USA.LFWA64TT.STSA.Q")) %>%
  mutate(series_name=case_when(str_detect(series_code,"RD-") ~ paste("Deflator,",s
eries_name),
                               str_detect(series_code,"RX-") ~ paste("Real,",serie
s_name),
                               str_detect(series_code,"Q.US.N.A.M.XDC.A") ~ paste
("Loans to non-financial corporations,",series_name),
                               str_detect(series_code,"Q.US.H.A.M.XDC.A") ~ paste
("Loans to households and NPISHs,",series_name),
                               str_detect(series_code,"Q.US.N.628") ~ paste("Prope
rty prices,",series_name),
                               TRUE ~ series_name)) %>%
  select(var_name=series_name,
         var_code=series_code,
         value,
         period)

df %<>%
  mutate(var_code=
           case_when(var_code=="A191RX-Q" ~ "gdp",
                     var_code=="A006RX-Q" ~ "inves",
                     var_code=="A191RD-Q" ~ "defgdp",
                     var_code=="A006RD-Q" ~ "definves",
                     var_code=="Q.US.H.A.M.XDC.A" ~ "loans_hh",
                     var_code=="Q.US.N.A.M.XDC.A" ~ "loans_nfc",
                     var_code=="Q.US.N.628" ~ "houseprice",
                     var_code=="PRS85006033" ~ "hours",
                     var_code=="PRS85006103" ~ "wage",
                     var_code=="USA.LFWA64TT.STSA.Q" ~ "pop",
                     var_code=="USA.IRLTLT01.ST.Q" ~ "longrate"))
```

We create a quarterly series of short-term interest rate taking the mean of the monthly series.

```
shortrate <-
  rdb("FED","H15",mask="129.FF.O") %>%
  mutate(period=paste(year(period),quarter(period),sep="-")) %>%
  group_by(period) %>%
  summarise(value=mean(value)) %>%
  mutate(var_code="shortrate",
         var_name="Monthly - Federal funds - Overnight",
         period=yq(period))
```

Special case of consumption

Private consumption is a special case : the series of consumption in level provided by the Bureau of Economic Analysis do not exist before 2002, we need to use the series of growth rates available before 2002 to deduce past consumption levels. As in Christiano et al. (2014), aggregate consumption is the sum of Non Durable Goods and Services. Durable Goods are associated with Investment. The following figure shows the three deduced series of consumption.

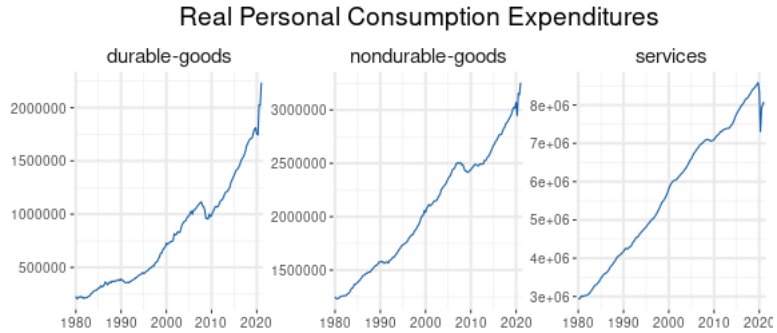
```
conso_level <-
  rdb(ids=c("BEA/NIPA-T20306/DDURRX-Q",
            "BEA/NIPA-T20306/DNDGRX-Q",
            "BEA/NIPA-T20306/DSERRX-Q")) %>%
  select(period,
         value,
         var_name=concept)

conso_rate <-
  rdb(ids=c("BEA/NIPA-T20301/DDURRL-Q",
            "BEA/NIPA-T20301/DNDGRL-Q",
            "BEA/NIPA-T20301/DSERRL-Q")) %>%
  select(period,
         value,
         var_name=concept)

conso_level_02 <-
  conso_level %>%
  filter(period=="2002-01-01")

conso <-
  conso_rate %>%
  filter(period <= "2002-01-01") %>%
  full_join(conso_level_02,by="var_name") %>%
  group_by(var_name) %>%
  arrange(desc(period.x)) %>%
  mutate(value = value.y / lag(cumprod((1 + value.x/100)^(1/4)))) %>%
  ungroup() %>%
  transmute(period=period.x,
            var_name,
            value) %>%
  na.omit() %>%
  bind_rows(conso_level) %>%
  filter(period >= "1980-01-01")

ggplot(conso,aes(period,value))+
  geom_line(colour=blueObsMacro)+
  facet_wrap(~var_name,ncol=3,scales = "free_y")+
  scale_x_date(expand = c(0.01,0.01)) +
  theme + xlab(NULL) + ylab(NULL)+
  ggtitle("Real Personal Consumption Expenditures")
```




```
conso %<>%
  mutate(
    var_code=case_when(
      var_name=="durable-goods" ~ "conso_d",
      var_name=="nondurable-goods" ~ "conso_nd",
      var_name=="services" ~ "conso_s"
    ),
    var_name=paste("Real Personal Consumption Expenditures,",var_name))
```

Financial data from FRED

```
fred_data <-
  bind_rows(data.frame(fred$series.observations(series_id = 'BAA',
                                                frequency="q",
                                                aggregation='avg'),
                       var_name="Moody's Seasoned Baa Corporate Bond Yield",
                       var_code="riskrate"),
            data.frame(fred$series.observations(series_id = 'WILL5000IND',
                                                frequency="q",
                                                aggregation='avg'),
                       var_name="Wilshire 5000 Total Market Index",
                       var_code="networth")) %>%
  transmute(period=ymd(date),
            value=as.numeric(value),
            var_code,var_name)
```

Final database and normalization

```
rawdata <-
  bind_rows(conso,df,shortrate,fred_data) %>%
  filter(year(period) >= 1980)
```

```
var_names <- unique(rawdata$var_name)
var_names <- gsub("Expenditures,.*","",var_names) %>% unique()
```

We can check the last date available for each variable.

```
maxDate <-
  rawdata %>%
  filter(var_code!="houseprice") %>%
  group_by(var_code) %>%
  summarize(maxdate=max(period)) %>%
  arrange(maxdate)
kable(maxDate)
```

var_code	maxdate
loans_hh	2020-10-01
loans_nfc	2020-10-01
conso_d	2021-01-01
conso_nd	2021-01-01
conso_s	2021-01-01
defgdp	2021-01-01
definves	2021-01-01
gdp	2021-01-01
hours	2021-01-01
inves	2021-01-01
longrate	2021-01-01
pop	2021-01-01
wage	2021-01-01
riskrate	2021-04-01
shortrate	2021-04-01
networth	2021-07-01

```
minmaxDate <- min(maxDate$maxdate)
rawdata %<>% filter(period <= minmaxDate) %>% select(-var_name)
rawdata %>%
  spread(var_code,value) %>%
  write.csv(file = "US_CMR_rawdata.csv", row.names= FALSE)
```

So we filter the database until 2020 Q4. Raw data can be downloaded directly here (http://shiny.cepremap.fr/data/US_CMR_rawdata.csv).

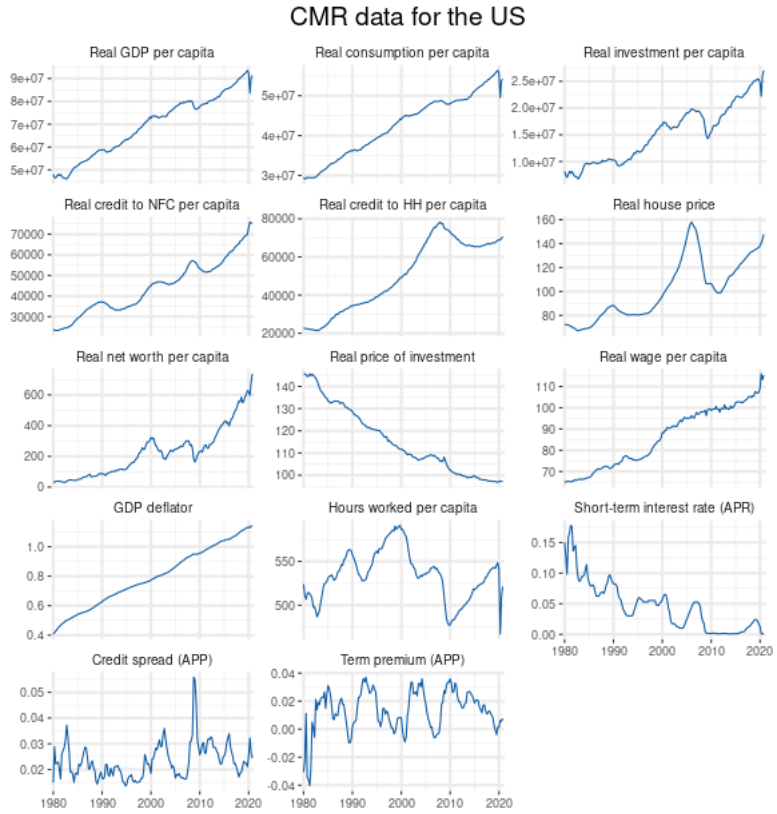
Then data are normalized by capita and price if needed. Eventually we have the 14 series : the 12 series similar to CMR plus loans to households and house price series.

```
US_CMR_data <-
  rawdata %>%
  spread(var_code,value) %>%
  transmute(period,
            gdp_rpc=1e+9*gdp/(1000*pop),
            conso_rpc=1e+9*(conso_nd+conso_s)/(1000*pop),
            inves_rpc=1e+9*(inves+conso_d)/(1000*pop),
            defgdp = defgdp/100,
            wage_rph=wage/defgdp,
            hours_pc=1e+9*hours/(1000*pop),
            pinves_defl=definves/defgdp,
            loans_nfc_rpc=1e+9*loans_nfc/(1000*pop)/defgdp,
            loans_hh_rpc=1e+9*loans_hh/(1000*pop)/defgdp,
            houseprice_defl=houseprice/defgdp,
            networth_rpc=1e+9*networth/(1000*pop)/defgdp,
            re=shortrate/100,
            slope=(longrate - shortrate)/100,
            creditspread=(riskrate - longrate)/100)

US_CMR_data %>%
  mutate(period=gsub(" ","",as.yearqtr(as.Date(period)))) %>%
  write.csv("US_CMR_data.csv", row.names=FALSE)
```

```
plot_US_CMR_data <-
  US_CMR_data %>%
  gather(var, value, - period) %>%
  mutate(var = as.factor(var))
levels(plot_US_CMR_data$var)<-listVar

ggplot(plot_US_CMR_data,aes(period,value))+
  geom_line(colour=blueObsMacro)+
  facet_wrap(~var,ncol=3,scales = "free_y")+
  scale_x_date(expand = c(0.01,0.01)) +
  theme + xlab(NULL) + ylab(NULL)+
  theme(strip.text=element_text(size=10),
        axis.text=element_text(size=9))+
  ggtitle("CMR data for the US")
```



You can also download ready-to-use (normalized) data for the estimation here (http://shiny.cepremap.fr/data/US_CMR_data.csv).

Bibliography

L Christiano, Roberto Motto, and Massimo Rostagno. Risk shocks. *American Economic Review*, 104(1):27–65, 2014. URL: http://www.aeaweb.org/articles.php?doi=10.1257/aer.104.1.27 (http://www.aeaweb.org/articles.php?doi=10.1257/aer.104.1.27), doi:10.1257/aer.104.1.27 (https://doi.org/10.1257/aer.104.1.27). ↩ 1 2

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL: https://www.R-project.org (https://www.R-project.org). ↩

RStudio Team. *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA, 2016. URL: http://www.rstudio.com/ (http://www.rstudio.com/). ↩

RELATED POSTS

Macroeconomic data for France, Germany, Italy, Spain & the Euro Area (/article/2021-02/five-countries-data/)

Description step by step to build automatic update of a macroeconomic database for France, Germany, Italy, Spain & the Euro Area.

MORE (/ARTICLE/2021-02/FIVE-COUNTRIES-DATA/)

Automating update of an international database for the Euro Area (/article/2019-12/open-EA-data/)

Description step by step to build automatic update of an international database for the Euro Area.

MORE (/ARTICLE/2019-12/OPEN-EA-DATA/)

Automating update of a fiscal database for the Euro Area (/article/2019-11/fipu-EA-data/)

Description step by step to build automatic update of a quarterly fiscal database for the Euro Area, similar to the Paredes, Pedregal...

MORE (/ARTICLE/2019-11/FIPU-EA-DATA/)

Automating update of a finance database for the Euro Area (/article/2016-06/cmr14-EA-data/)

Description step by step to build automatic update a finance database for the Euro Area, similar to the Christiano, Motto and Rostagno...

MORE (/ARTICLE/2016-06/CMR14-EA-DATA/)

Follow us

Gitlab (https://git.nomics.world/macro)

Twitter (https://twitter.com/obsmacro)

Links

Dynare (http://www.dynare.org/)

DBnomics (https://db.nomics.world/)

R-bloggers (http://www.r-bloggers.com/)