Could a large scale asset purchase programme have mitigated the Great Depression

Garabedian, Garo∗,Γ - Stuart, Rebecca∗

∗Central Bank of Ireland, ΓGhent University

ASSA meeting, Atlanta - January 5th, 2019
Outline

- Recent financial crisis: unconventional monetary policy measures
  - Large scale asset purchase programs
  - Sharp interest rate ↓ (towards effective ZLB)
- **What would have been the impact of more aggressive monetary policy during the Great Depression?**
  - Bayesian VAR to estimate forecasts conditional on alternative policy paths
- Contrast with recent policy:
  - Gold standard
    - Temin (1976): Fed limited in what it could do
  - No announcement
    - No explicit forward guidance
Historical Context

- **Banking Act 1932**
  - Changed proportion of gold required to back notes
  - Thus freeing gold reserves

- **February 1932**
  - OMPC authorised purchases of government securities

- **July 9, Gov. McDougal of Chicago Fed:**
  - “...we believe that the additional purchases made were much too large and have resulted in creating abnormally low rates for short-term US Government securities.”
Macroeconomy

- Highlighted bars: Purchase period (April-July 1932), 'Roosevelt bank holiday' (March/April 1933)
- 'Double Bottom' (Burns and Mitchell, 1946)
Fed’s role during Great Depression has been criticised:

- Little to mitigate effects of crisis
- No prevailing wisdom about how to respond to downturn
- Although *official rates* were reduced at start of crisis
  - Never reached ZLB
  - Even were raised in late 1931 and early 1932 in response to strong outflows of gold after Britain abandoned gold standard
- Government bond *purchase program*: limited in time

Instead, *Roosevelt’s ‘bank holiday’* in March/April 1933 and effective devaluation of dollar
Data

- Monthly data: 1919M1-1934M7

- Data set tries to mimic what Fed would have looked at as most likely to influence policy (Iversen et al., 2014)
  - Federal Reserve Bulletins: National Summary of Business Conditions

- Data sources: Fred, Alfred, Fraser, NBER historical database, Shiller
Data Categories

1. **Prices**
   - CPI, PPI, wholesale price fuel and lighting

2. **Business cycle**
   - Industrial production, department store sales

3. **Labour market**
   - Factory employment, factory earnings

4. **Financial variables**
   - Money stock, S&P composite stock price index, yield spread (10y - 3m), exchange rate with Swiss Franc

5. **Stress measures**
   - Liabilities of commercial business failures, spread (Baa-rated corporate bonds and LT govt bonds), spread between secured and unsecured money market rates

6. **Monetary policy variables**
   - Fed purchases of government securities, NY Fed discount rate
Large Bayesian VAR Model

- **Large Bayesian VAR**
  - 17 variables, 12 lags
  - Core variables + labour, financial, external and monetary variables
    - Gambacorta et al. (2014): need to capture spillovers between real economy and financial markets

- Gains from large Bayesian VARs (De Mol et al., 2008):
  - Estimation through shrinkage of parameters

- **Dummy observation prior** (Banbura et al., 2010)
  - Allows us to match Minnesota moments and integrate sum of coefficients
    - Consistent with unit root or cointegration processes
Large Bayesian VAR Model

- Desire to include many macro variables (Gambacorta et al., 2014)
  → “Curse of dimensionality”
  - Uncertainty, financial turmoil and economic risk variables to unravel exogenous changes in CB balance sheet from endogenous intervention

- Quick proliferation of parameters that have to be reliably estimated in large dimensional systems
  - Trade-off between
    - Misspecification and forecast accuracy
    - Issues of collinearity and overfitting

- High number of parameters cannot be well estimated by ML/OLS,
  - Recent developments in macroeconometrics → 2 approaches able to deal with this complexity:
    - **Bayesian VARs** and dynamic factor models (Banbura et al., 2014)
Scenario Analysis

- **Conditional forecasting methodology** (Waggoner and Zha, 1999)
  - Many policy applications:

- Government securities purchased April - July 1932: $121 million to $399 million per month
  - Our path: 12 more months at $220 million per month (Aug 1932 to July 1933)
    - Taper for 6 months until just $105 million of purchases are made in January 1934
  - Reduction in NY Fed discount rate from 2 percent (Aug 1932) to 0.25 percent in (Aug 1933)
Path 1: NY Fed Discount Rate

![Graph showing the NY Fed Discount Rate and QE Path from 1929 to 1934.](image)
Path 2: Fed Purchases of Gov Securities

Historical Purchases
Relative Size of Purchases
Consumer Price Inflation

![Graph showing Consumer Price Inflation from 1928 to 1934](image)

- CPI
- Median
- Lower
- Upper
Industrial Production Growth

- IP Growth
- Median
- Lower
- Upper

Years:
- 1929
- 1930
- 1931
- 1932
- 1933
- 1934

Values:
- 140
- 120
- 100
- 80
- 60
- 40
- 20
- 0
- -20
- -40
- -60
Money Stock Growth
Exchange Rate
To test whether extending the purchase programme really mattered, we analyse several alternative scenarios:

1. No asset purchases, only the prior interest rate cut
2. Smaller asset purchase programme, and the prior rate cut:
   - $180 million per month for 12 months (previously $220)
   - Tapering for 6 months: $65 million in the final month
CPI Inflation, Alternative Scenario

[Graph showing CPI inflation over time with three lines representing different scenarios: Alternative - no purchase programme, Baseline scenario, and Alternative - lower purchases.]
IP Growth, Alternative Scenario

![Graph showing IP growth over time with different scenarios]

- **Alternative - no purchase programme**
- **Baseline scenario**
- **Alternative - lower purchases**

**Years:** 1932-08, 1932-12, 1933-04, 1933-08, 1933-12
Conclusion

- More expansionary monetary policy could have eased the Great Depression
  - Positive growth in prices and output sooner
- Federal Reserve could have significantly improved economic outcomes
  - Both purchases of government securities as well as the interest rate drop are instrumental
  - But would have increased the money supply substantially
- Large impact on the exchange rate
Minnesota Prior

- Minnesota prior for coefficients can be retrieved by setting following moments (Blake and Mumtaz, 2012)

\[
E \left[ (A_k)_{ij} \right] = \begin{cases} 
\delta_i & j = i, k = 1 \\
0 & \text{otherwise}
\end{cases}, \quad V \left[ (A_k)_{ij} \right] = \\
\begin{cases} 
\frac{\lambda^2}{k^2} & j = i \\
\vartheta \frac{\lambda^2}{k^2} \frac{\sigma_i^2}{\sigma_j^2} & \text{otherwise}
\end{cases}
\]  

- Coefficients \(A_1, \ldots, A_p\) considered independent and normally distributed
- \(\delta_i\) : prior coefficient mean for first lag of dependent variables
Minnesota Prior

- Hyperparameter $\lambda$: general tightness of prior distribution around random walk or white noise component
- $1/k^2$: rate at which prior variance decreases when lag length increases
- Ratio $\frac{\sigma_i^2}{\sigma_j^2}$ takes into account differences in scale of data
- Coefficient $\vartheta \in (0, 1)$ characterizes the extent to which lags of other variables are ‘less important’ than own lags
- Covariance matrix of residuals is considered to be diagonal, fixed and known
  - $\Psi = \Sigma$, with $\Sigma = \text{diag} \left( \sigma_1^2, \ldots, \sigma_n^2 \right)$
- Prior on intercept is diffuse
Following Banbura et al (2010), we implement dummy observation prior by appending $T_d$ dummy observations, as expressed in $Y_d$ and $X_d$, to the system:

Where coefficients have normal prior and covariance matrix has normal inverted Wishart prior:

$vec(B) \mid \Psi \sim N(vec(B_0), \Psi \otimes \Omega_0)$, $\Psi \sim iW(S_0, \alpha_0)$.

$$Y_d = \begin{pmatrix}
\text{diag}(\delta_1 \sigma_1, \ldots, \delta_n \sigma_n) / \lambda \\
0_{n(p-1) \times n}
\end{pmatrix}$$

$$X_d = 
\begin{pmatrix}
J_p \otimes \text{diag}(\delta_1 \sigma_1, \ldots, \delta_n \sigma_n) / \lambda & 0_{np \times 1} \\
0_{n \times np} & 0_{n \times 1}
\end{pmatrix}
$$

with $J_p = \text{diag}(1, 2 \ldots, p)$. 

(3)
Dummy Observation Prior

- Different segments:
  - First block of dummies represents the prior beliefs on the AR coefficients
  - Second block summarizes prior for the covariance matrix
  - Third block describes the uninformative prior for the intercept

- We retrieve required structures $Y^*$ and $X^*$ by adding dummies $Y_d$ and $X_d$ to the original data:

$$Y^* = [Y, Y_d], \quad X^* = [X, X_d] \quad (4)$$

- Using this appended data, the conditional distributions can be integrated in the Gibbs sampling algorithm
  - Results are based on 15000 draws from the Gibbs sampler, with a burn-in of 10000
Conditional Forecasting

- "Conditional-on-observables" (Banbura et al., 2015)
  - Outcome for macro-financial variables, given path for policy rate and purchases
  - What would have happened if Fed continued purchases at every point
    - Hsieh and Romer (2006)
- Alternatively, what would happen if there was a series of MP surprises at each point
  - "Structural scenario analysis"
  - (Antolin-Diaz et al., 2018)
Conditional Forecasting

Consider a VAR(1) model (Blake and Mumtaz, 2014):

$$Y_t = c + BY_t + A_0 \varepsilon_t \quad (5)$$

with $Y_t$ representing a $T \times N$ matrix of endogenous variables

- $\varepsilon_t$ denoting the uncorrelated structural shocks
- $A_0 A_0' = \Sigma$. 
  - $\Sigma$ represents the variance of the reduced VAR residuals

When we iterate equation (5) $K$ times forward, we retrieve

$$Y_{t+K} = c \sum_{j=0}^{K} B^j + B^j Y_{t-1} + A_0 \sum_{j=0}^{K} B^j \varepsilon_{t+K-j} \quad (6)$$
Conditional Forecasting

- Hence, when we place restrictions on future path of $j^{th}$ variable in $Y_t$, this also induces restrictions on other variables in system
  - If we re-structure equation (6) this becomes more visible:

\[
Y_{t+K} - c \sum_{j=0}^{K} B^j - B^j Y_{t-1} = A_0 \sum_{j=0}^{K} B^j \varepsilon_{t+K-j} \quad (7)
\]

- When we constrain some of variables in dataset to fixed path, this means that future innovations on right hand side of equation will have restrictions as well.
  - These constraints on future innovations are defined in Waggoner and Zha (1999) as:

\[
R \varepsilon = r \quad (8)
\]
Elements of $r$ contain path for constrained variables minus unconditional forecasts of constrained variables.

- Elements of matrix $R$ are impulse responses of constrained variables to structural shocks $\varepsilon$ over desired forecasting horizon.
- A least square solution for constrained shocks in (8) is given by Doan et al. (1983):

$$\varepsilon = R' (R'R)^{-1} r \ (9)$$

Inserting these constrained innovations in equation (5) allow us to calculate conditional forecasts.
Historical Gov Securities Purchases
Real Wage Growth
Food Price Inflation