Dynamic simulation

July 23, 2025

1 Dynamic system

Suppose we need to predict the DISTRIBUTION of balance sheet items at t whose value at $t-1$ is known. The process is decomposed into applying accounting identities and prediction of flow variables.

Key takeaways:

1. Balance sheet items are stock variables, and all follow accounting identities.

2. Items in income statements and the statement of financial condition (cash flow statements) are flow variables

3. Items in cash flow statements follow accounting identities

4. Some variables in income statements are decision variables, and are assumed to follow specific function forms, and are estimated.

5. Some variables in income statements are not decision variables but are also assumed to follow specific function forms and estimated, for example, financial income and financial expenses.

6. The rest of the variables in income statements follows accounting identities, for example, net income.

1.1 Variable classification

A first classification is summarized in the table below.

Follow Accounting Identities	Assumed and Estimated
Balance sheet items: MA_t, BU_t, OFA_t, CA_t, SC_t, RR_t, OUR_t, CMA_t, ASD_t, PF_t^{t-5}, PF_t^{t-4}, PF_t^{t-3}, PF_t^{t-2}, PF_t^{t-1}, PF_t, LL_t, CL_t, URE_t	Decision variables in income statements: I_t^{MA}, I_t^{BU}, $dofa_t$, dca_t, dsc_t, drr_t, $dour_t$, p_t^{allo}, $TDEP_t^{MA}$, zpf_t, dll_t, dcl_t
Cash flow items: $cashfl_t$, $mcash_t$	Other income statements items: $EDEP_t^{MA}$,S_t^{MA}, $EDEP_t^{BU}$, $OIBD_t$,FI_t, FE_t, GC_t, OA_t, TL_t, OTA_t, $TDEP_t^{BU}$, ROT_t
Some Income statement items: EBT_t, TA_t, TAX_t, $FTAX_t$, NBI_t, OL_t	

1.2 List of Variables

Appendix A of paper 04.

1.3 The system Dynamics for the firm

The dynamics are:

$$
\begin{aligned}
MA_t &= MA_{t-1} + I_t^{MA} - S_t^{MA} - EDEP_t^{MA} \\
BU_t &= BU_{t-1} + I_t^{BU} - EDEP_t^{BU} \\
OFA_t &= OFA_{t-1} + dofa_t \\
CA_t &= CA_{t-1} + dca_t \\
SC_t &= SC_{t-1} + dsc_t \\
RR_t &= RR_{t-1} + drr_t \\
OUR_t &= OUR_{t-1} + dour_t \\
ASD_t &= ASD_{t-1} + (TDEP_t^{MA} - EDEP_t^{MA}) \\
CMA_t &= MA_t - ASD_t \\
p_t^{allo} &= \max[0, \min(p_t^{allo}, \eta \times pbase_t)] \\
MPA_t &= \max[0, \eta \times pbase_t] \\
dmpa_t &= MPA_t - p_t^{allo} \\
ddmpa_t &= dmpa_t - dmpa_{t-1} \\
PF_t &= PF_{t-1} + p_t^{allo} - zpf_t \\
LL_t &= LL_{t-1} + dll_t \\
CL_t &= CL_{t-1} + dcl_t \\
EBT_t &= OIBD_t - EDEP_t^{BU} + FI_t - FE_t - TDEP_t^{MA} - p_t^{allo} + zpf_t + OA_t \\
TA_t &= OTA_t - TDEP_t^{BU} - OL_{t-1} \\
TAX_t &= \tau \max[0, (EBT_t - TL_t + TA_t)] \\
FTAX_t &= TAX_t - ROT_t \\
NBI_t &= EBT_t - FTAX_t \\
OL_t &= \text{abs}\,(\min[0, (EBT_t - TL_t + TA_t)]) + OL_{t-1} \\
cashfl_t &= OIBD_t + FI_t - FE_t + OA_t - FTAX_t - DIV_{t-1} + dsc_t + dcl_t \\
&\quad + dll_t + dour_t - I_t^{MA} + S_t^{MA} - I_t^{BU} - dofa_t - dca_t \\
URE_t &= URE_{t-1} + NBI_t - DIV_{t-1} - drr_t - cashfl_t \\
mcash_t &= URE_{t-1} + NBI_t - drr_t \\
MTDM_t &= 2.53/2.54 \\
dmcash_t &= mcash_t - cashfl_t \\
ddmcash_t &= dmcash_t - dmcash_{t-1} \\
dmtdm_t &= MTDM_t - TDEP_t^{MA} \\
ddmtdm_t &= dmtdm_t - dmtdm_{t-1} \\
DIV_t &= \max[0, \min(cashfl_t, mcash_t)] \\
dgnp &= 0 \\
realr &= 0
\end{aligned}
$$

1.4 Behavior

The behaviors marked in red are divided into several groups for clarity:

1.4.1 MA and BU Related Variables

$$\begin{aligned} EDEP_t^{MA} = {} & \gamma_0^{EMA} + \gamma_2^{EMA} I_{t-1}^{MA} + \gamma_3^{EMA} dca_{t-1} + \gamma_4^{EMA} dcashfl_{t-1} + \gamma_5^{EMA} ddmpa_{t-1} \\ & + \gamma_6^{EMA} ddmcash_{t-1} + \gamma_7^{EMA} ddmtdm_{t-1} + \gamma_8^{EMA} TDEP_{t-1}^{MA} + \gamma_{10}^{EMA} EDEP_{t-1}^{BU} \\ & + \gamma_{14}^{EMA} MA_{t-1} + \gamma_{32}^{EMA} dcl_{t-1} \end{aligned}$$

$$\begin{aligned} S_t^{MA} = {} & \gamma_0^{SMA} + \gamma_2^{SMA} I_{t-1}^{MA} + \gamma_3^{SMA} dca_{t-1} + \gamma_4^{SMA} dcashfl_{t-1} + \gamma_5^{SMA} ddmpa_{t-1} \\ & + \gamma_6^{SMA} ddmcash_{t-1} + \gamma_7^{SMA} ddmtdm_{t-1} + \gamma_8^{SMA} TDEP_{t-1}^{MA} + \gamma_9^{SMA} EDEP_t^{MA} \\ & + \gamma_{10}^{SMA} EDEP_{t-1}^{BU} + \gamma_{14}^{SMA} MA_{t-1} + \gamma_{32}^{SMA} dcl_{t-1} \end{aligned}$$

$$\begin{aligned} I_t^{MA} = {} & \gamma_0^{IMA} + \gamma_2^{IMA} I_{t-1}^{MA} + \gamma_3^{IMA} dca_{t-1} + \gamma_4^{IMA} dcashfl_{t-1} + \gamma_5^{IMA} ddmpa_{t-1} + \gamma_6^{IMA} ddmcash_{t-1} \\ & + \gamma_7^{IMA} ddmtdm_{t-1} + \gamma_8^{IMA} TDEP_{t-1}^{MA} + \gamma_9^{IMA} EDEP_t^{MA} + \gamma_{10}^{IMA} EDEP_{t-1}^{BU} + \gamma_{11}^{IMA} S_t^{MA} \\ & + \gamma_{32}^{IMA} dcl_{t-1} \end{aligned}$$

$$\begin{aligned} EDEP_t^{BU} = {} & \gamma_0^{EBU} + \gamma_1^{EBU} I_t^{MA} + \gamma_3^{EBU} dca_{t-1} + \gamma_4^{EBU} dcashfl_{t-1} + \gamma_5^{EBU} ddmpa_{t-1} + \gamma_6^{EBU} ddmcash_{t-1} \\ & + \gamma_9^{EBU} EDEP_t^{MA} + \gamma_{11}^{EBU} S_t^{MA} + \gamma_{12}^{EBU} BU_{t-1} + \gamma_{32}^{EBU} dcl_{t-1} \end{aligned}$$

$$\begin{aligned} I_t^{BU} = {} & \gamma_0^{IBU} + \gamma_1^{IBU} I_t^{MA} + \gamma_3^{IBU} dca_{t-1} + \gamma_4^{IBU} dcashfl_{t-1} + \gamma_5^{IBU} ddmpa_{t-1} + \gamma_6^{IBU} ddmcash_{t-1} \\ & + \gamma_9^{IBU} EDEP_t^{MA} + \gamma_{10}^{IBU} EDEP_t^{BU} + \gamma_{11}^{IBU} S_t^{MA} + \gamma_{32}^{IBU} dcl_{t-1} \end{aligned}$$

$$\begin{aligned} TDEP_t^{MA} = {} & \gamma_0^{TMA} + \gamma_1^{TMA} I_t^{MA} + \gamma_4^{TMA} dcashfl_{t-1} + \gamma_5^{TMA} ddmpa_{t-1} + \gamma_6^{TMA} ddmcash_{t-1} + \gamma_7^{TMA} EDEP_t^{MA} \\ & + \gamma_{11}^{TMA} S_t^{MA} \end{aligned}$$

$$\begin{aligned} TDEP_t^{BU} = {} & \gamma_0^{TBU} + \gamma_1^{TBU} I_t^{MA} + \gamma_3^{TBU} dca_t + \gamma_4^{TBU} dcashfl_{t-1} + \gamma_5^{TBU} ddmpa_{t-1} \\ & + \gamma_6^{TBU} ddmcash_{t-1} + \gamma_9^{TBU} EDEP_t^{MA} + \gamma_{11}^{TBU} S_t^{MA} + \gamma_{15}^{TBU} BU_{t-1} + \gamma_{33}^{TBU} dcl_t \end{aligned}$$

1.4.2 Current Assets and Liabilities

$$dofa_t = \gamma_0^{OFA} + \gamma_4^{OFA} dcashfl_{t-1} + \gamma_5^{OFA} ddmpa_{t-1} + \gamma_6^{OFA} ddmcash_{t-1}$$

$$\begin{aligned} dca_t = {} & \gamma_0^{CA} + \gamma_1^{CA} I_t^{MA} + \gamma_2^{CA} I_t^{BU} + \gamma_3^{CA} dcashfl_{t-1} + \gamma_5^{CA} ddmpa_{t-1} + \gamma_6^{CA} ddmcash_{t-1} \\ & + \gamma_9^{CA} EDEP_t^{MA} + \gamma_{10}^{CA} EDEP_t^{BU} + \gamma_{11}^{CA} S_t^{MA} + \gamma_{32}^{CA} dcl_{t-1} \end{aligned}$$

$$dll_t = \gamma_0^{LL} + \gamma_4^{LL} dcashfl_{t-1} + \gamma_5^{LL} ddmpa_{t-1} + \gamma_6^{LL} ddmcash_{t-1}$$

$$\begin{aligned} dcl_t = {} & \gamma_0^{CL} + \gamma_1^{CL} I_t^{MA} + \gamma_2^{CL} I_t^{BU} + \gamma_3^{CL} dca_t + \gamma_4^{CL} dcashfl_{t-1} + \gamma_5^{CL} ddmpa_{t-1} + \gamma_6^{CL} ddmcash_{t-1} \\ & + \gamma_9^{CL} EDEP_t^{MA} + \gamma_{10}^{CL} EDEP_t^{BU} + \gamma_{11}^{CL} S_t^{MA} \end{aligned}$$

$$dsc_t = \gamma_0^{SC} + \gamma_4^{SC} dcashfl_{t-1} + \gamma_6^{SC} ddmcash_{t-1}$$

$$drr_t = \gamma_0^{RR} + \gamma_6^{RR} ddmcash_{t-1}$$

1.4.3 Income Statement Variables

$$\begin{aligned}OIBD_t = \gamma_0^{OIBD} &+ \gamma_1^{OIBD} I_t^{MA} + \gamma_2^{OIBD} I_t^{BU} + \gamma_3^{OIBD} dca_t + \gamma_4^{OIBD} dcashfl_{t-1} + \gamma_5^{OIBD} ddmpa_{t-1} \\ &+ \gamma_9^{OIBD} EDEP_t^{MA} + \gamma_{10}^{OIBD} EDEP_t^{BU} + \gamma_{11}^{OIBD} S_t^{MA} + \gamma_{14}^{OIBD} MA_{t-1} + \gamma_{15}^{OIBD} BU_{t-1} \\ &+ \gamma_{16}^{OIBD} CA_{t-1} + \gamma_{32}^{OIBD} CL_{t-1} + \gamma_{33}^{OIBD} dcl_t\end{aligned}$$

$$\begin{aligned}FI_t = \gamma_0^{FI} &+ \gamma_1^{FI} I_t^{MA} + \gamma_2^{FI} I_t^{BU} + \gamma_3^{FI} dca_t + \gamma_9^{FI} EDEP_t^{MA} + \gamma_{10}^{FI} EDEP_t^{BU} + \gamma_{11}^{FI} S_t^{MA} \\ &+ \gamma_{14}^{FI} MA_{t-1} + \gamma_{15}^{FI} BU_{t-1} + \gamma_{16}^{FI} CA_{t-1} + \gamma_{17}^{FI} OFA_{t-1} + \gamma_{31}^{FI} dofa_t\end{aligned}$$

$$\begin{aligned}FE_t = \gamma_0^{FE} &+ \gamma_1^{FE} I_t^{MA} + \gamma_2^{FE} I_t^{BU} + \gamma_3^{FE} dca_t + \gamma_9^{FE} EDEP_t^{MA} + \gamma_{10}^{FE} EDEP_t^{BU} + \gamma_{11}^{FE} S_t^{MA} \\ &+ \gamma_{14}^{FE} MA_{t-1} + \gamma_{15}^{FE} BU_{t-1} + \gamma_{16}^{FE} CA_{t-1} + \gamma_{17}^{FE} OFA_{t-1} + \gamma_{18}^{FE} CL_{t-1} + \gamma_{19}^{FE} dcl_t \\ &+ \gamma_{20}^{FE} LL_{t-1} + \gamma_{21}^{FE} dll_t + \gamma_{32}^{FE} dofa_t\end{aligned}$$

1.4.4 Other Financial Variables

$$zpf_t = \gamma_0^{ZPF} + \gamma_4^{ZPF} dcashfl_{t-1} + \gamma_5^{ZPF} ddmpa_{t-1} + \gamma_6^{ZPF} ddmcash_{t-1} + \gamma_{13}^{ZPF} p_{t-1}^{Allo}$$

$$dour_t = \gamma_0^{OUR} + \gamma_4^{OUR} dcashfl_{t-1} + \gamma_5^{OUR} ddmpa_{t-1} + \gamma_6^{OUR} ddmcash_{t-1}$$

$$\begin{aligned}GC_t = \gamma_0^{GC} &+ \gamma_{10}^{GC} EDEP_t^{BU} + \gamma_{22}^{GC} OIBD_t + \gamma_{23}^{GC} FI_t + \gamma_{24}^{GC} FE_t + \gamma_{25}^{GC} TDEP_t^{MA} + \gamma_{26}^{GC} zpf_t \\ &+ \gamma_{27}^{GC} dour_t\end{aligned}$$

$$OA_t = \gamma_0^{OA} + \gamma_{28}^{OA} dour_t + \gamma_{29}^{OA} GC_t$$

$$p_t^{allo} = \hat{\gamma}_0^{PPF} + \hat{\gamma}_4^{PPF} dcashfl_{t-1} + \hat{\gamma}_{33}^{PPF} MPA_t + \hat{\gamma}_{34}^{PPF} dmpa_{t-1} + \hat{\gamma}_6^{PPF} ddmcash_{t-1} + \hat{\gamma}_{12}^{PPF} zpf_t$$

1.4.5 Tax Related Variables

$$\begin{aligned}TL_t = \gamma_0^{TL} &+ \gamma_{10}^{TL} EDEP_t^{BU} + \gamma_{13}^{TL} \boldsymbol{p_{t-1}^{allo}} + \gamma_{22}^{TL} OIBD_t + \gamma_{23}^{TL} FI_t + \gamma_{24}^{TL} FE_t + \gamma_{25}^{TL} TDEP_t^{MA} \\ &+ \gamma_{26}^{TL} zpf_t + \gamma_{27}^{TL} dour_t\end{aligned}$$

$$\begin{aligned}OTA_t = \gamma_0^{OTA} &+ \gamma_{10}^{OTA} EDEP_t^{BU} + \gamma_{13}^{OTA} \boldsymbol{p_{t-1}^{allo}} + \gamma_{22}^{OTA} OIBD_t + \gamma_{23}^{OTA} FI_t + \gamma_{24}^{OTA} FE_t \\ &+ \gamma_{25}^{OTA} TDEP_t^{MA} + \gamma_{26}^{OTA} zpf_t + \gamma_{27}^{OTA} dour_t + \gamma_{28}^{OTA} TL_t\end{aligned}$$

$$\begin{aligned}ROT_t = \gamma_0^{ROT} &+ \gamma_{10}^{ROT} EDEP_t^{BU} + \gamma_{13}^{ROT} p_t^{allo} + \gamma_{22}^{ROT} OIBD_t + \gamma_{23}^{ROT} FI_t + \gamma_{24}^{ROT} FE_t + \gamma_8^{ROT} TDEP_t^{MA} \\ &+ \gamma_{12}^{ROT} zpf_t + \gamma_{25}^{ROT} dour_t + \gamma_{26}^{ROT} TL_t + \gamma_{27}^{ROT} OTA_t + \gamma_{28}^{ROT} TDEP_t^{BU}\end{aligned}$$

2 Estimation

Parameters to be estimated are in section 1.4. To estimate these parameters with real life data, we also add control variables.

2.1 Control variables

Though not directly specified in Section 2 and 3 in 04, there are multiple control variables to be included in the dynamic system.

Four types of control variables: exogenous individual, exogenous macro, endogenous individual, financial ratios.

exogenous individual	exogenous macro	endogenous individual
maturity, location, concentration	GNP, real rate	DI_t^{MA}, DI_t^{BU}, $Ddofa_t$, $Dddl_t$, dsc_t, $DTDEP_t$, $Dzpf_t$, $Ddour_t$

Specifically, all the endogenous individual variables are dummy variables. The application of financial ratios are only exhibited in the appendix when viewing the results of the estimation.

2.2 Estimation methods

The second classification is for variables that are estimated. Two methods are used:

1. Two stage estimation. Estimate the probability of behavior first, then for those who take the behavior.
2. One stage estimation. Directly estimate the parameters.

2.3 The exact form of equations

Controls for each equation are different. Luckily, in appendix we can find details of which controls are used for each 24 estimation. A summary is listed below:

2.3.1 EDEPMA

First stage:

$$\begin{aligned}\eta_{EDEPMA} = {} & \beta_0 + \beta_1 ddmcash_{t-1} + \beta_2 dcash_{t-1} + \beta_3 TDEP_{t-1}^{MA} + \beta_4 MA_{t-1} + \beta_5 I_{t-1}^{MA} + \beta_6 (I_{t-1}^{MA})^2 \\ & + \beta_7 EDEP_{t-1}^{BU} + \beta_8 (EDEP_{t-1}^{BU})^2 + \beta_9 ddmtdm_{t-1} + \beta_{10}(ddmtdm_{t-1})^2 + \beta_{11} dca_{t-1} \\ & + \beta_{12} ddmpa_{t-1} + \beta_{13}(ddmpa_{t-1})^2 + \beta_{14} dcl_{t-1} + \beta_{15} dgnp_t + \beta_{16} FAAB_t \\ & + \beta_{17} Public_t + \beta_{18} ruralare_t + \beta_{19} largcity_t + \beta_{20} market_t + \beta_{21} marketw_t\end{aligned}$$

Second stage:

$$\begin{aligned}EDEP_t^{MA} = {} & \gamma_0 + \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 TDEP_{t-1}^{MA} + \gamma_4 MA_{t-1} + \gamma_5 I_{t-1}^{MA} + \gamma_6 (I_{t-1}^{MA})^2 \\ & + \gamma_7 EDEP_{t-1}^{BU} + \gamma_8 (EDEP_{t-1}^{BU})^2 + \gamma_9 ddmtdm_{t-1} + \gamma_{10}(ddmtdm_{t-1})^2 + \gamma_{11} dca_{t-1} \\ & + \gamma_{12} ddmpa_{t-1} + \gamma_{13}(ddmpa_{t-1})^2 + \gamma_{14} dcl_{t-1} + \gamma_{15} dgnp_t + \gamma_{16} FAAB_t \\ & + \gamma_{17} Public_t + \gamma_{18} ruralare_t + \gamma_{19} largcity_t + \gamma_{20} market_t + \gamma_{21} marketw_t\end{aligned}$$

2.3.2 SMA

First stage:

$$\begin{aligned}\eta = Intercept_1 + Intercept_2 + {} & \beta_{1,1} ddmcash_{t-1} + \beta_{2,1} dcash_{t-1} + \beta_{3,1} TDEP_{t-1}^{MA} + \beta_{4,1} EDEP_t^{MA} \\ & + \beta_{5,1}(EDEP_t^{MA})^2 + \beta_{6,1} MA_{t-1} + \beta_{7,1} I_{t-1}^{BU} + \beta_{8,1}(I_{t-1}^{BU})^2 \\ & + \beta_{9,1} EDEP_{t-1}^{BU} + \beta_{10,1}(EDEP_{t-1}^{BU})^2 + \beta_{11,1} ddmtdm_{t-1} + \beta_{12,1}(ddmtdm_{t-1})^2 \\ & + \beta_{13,1} dca_{t-1} + \beta_{14,1} ddmpa_{t-1} + \beta_{15,1}(ddmpa_{t-1})^2 + \beta_{16,1} dcl_{t-1} \\ & + \beta_{17,1} dgnp_t + \beta_{18,1} FAAB_t + \beta_{19,1} Public_t + \beta_{20,1} ruralare_t \\ & + \beta_{21,1} largcity_t + \beta_{22,1} market_t + \beta_{23,1} marketw_t\end{aligned}$$

Second stage for positive sales:

$$\begin{aligned}S_t^{MA} = {} & \gamma_0 + \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 EDEP_t^{MA} + \gamma_4 MA_{t-1} + \gamma_5 I_{t-1}^{BU} + \gamma_6 (I_{t-1}^{BU})^2 \\ & + \gamma_7 EDEP_{t-1}^{BU} + \gamma_8 (EDEP_{t-1}^{BU})^2 + \gamma_9 ddmtdm_{t-1} + \gamma_{10} dca_{t-1} + \gamma_{11} ddmpa_{t-1} \\ & + \gamma_{12}(ddmpa_{t-1})^2 + \gamma_{13} dcl_{t-1} + \gamma_{14}(dcl_{t-1})^2 + \gamma_{15} dgnp_t + \gamma_{16} FAAB_t \\ & + \gamma_{17} Public_t + \gamma_{18} ruralare_t + \gamma_{19} largcity_t + \gamma_{20} market_t + \gamma_{21} marketw_t\end{aligned}$$

Second stage for negative sales:

$$\begin{aligned}S_t^{MA} = \delta_0 &+ \delta_1 ddmcash_{t-1} + \delta_2 dcash_{t-1} + \delta_3 dcat_{t-1} + \delta_4 dcl_{t-1} + \delta_5 TDEP_{t-1}^{MA} + \delta_6 EDEP_t^{MA} \\ &+ \delta_7 (EDEP_t^{MA})^2 + \delta_8 MA_{t-1} + \delta_9 I_{t-1}^{BU} + \delta_{10} EDEP_{t-1}^{BU} + \delta_{11} (EDEP_{t-1}^{BU})^2 \\ &+ \delta_{12} ddmtdm_{t-1} + \delta_{13} ddmpa_{t-1} + \delta_{14} dgnp_t + \delta_{15} FAAB_t + \delta_{16} Public_t \\ &+ \delta_{17} ruralare_t + \delta_{18} largcity_t + \delta_{19} market_t + \delta_{20} marketw_t\end{aligned}$$

2.3.3 IMA

$$\begin{aligned}I_t^{MA*} = \beta_0 &+ \beta_1 ddmcash_{t-1} + \beta_2 dcash_{t-1} + \beta_3 S_t^{MA} + \beta_4 I_{t-1}^{BU} + \beta_5 EDEP_{t-1}^{BU} + \beta_6 (EDEP_{t-1}^{BU})^2 \\ &+ \beta_7 EDEP_t^{MA} + \beta_8 TDEP_{t-1}^{MA} + \beta_9 (TDEP_{t-1}^{MA})^2 + \beta_{10} ddmtdm_{t-1} + \beta_{11} dca_{t-1} \\ &+ \beta_{12} ddmpa_{t-1} + \beta_{13} (ddmpa_{t-1})^2 + \beta_{14} dcl_{t-1} + \beta_{15} dgnp_t + \beta_{16} FAAB_t \\ &+ \beta_{17} Public_t + \beta_{18} ruralare_t + \beta_{19} largcity_t + \beta_{20} market_t + \beta_{21} marketw_t \\ I_t^{MA} &= \max(0, I_t^{MA*})\end{aligned}$$

2.3.4 EDEPBU

First stage:

$$\begin{aligned}\eta_{EDEPBU} = \beta_0 &+ \beta_1 ddmcash_{t-1} + \beta_2 dcash_{t-1} + \beta_3 EDEP_t^{MA} + \beta_4 (EDEP_t^{MA})^2 + \beta_5 S_t^{MA} + \beta_6 I_t^{MA} \\ &+ \beta_7 BU_{t-1} + \beta_8 (BU_{t-1})^2 + \beta_9 dca_{t-1} + \beta_{10} ddmpa_{t-1} + \beta_{11} dcl_{t-1} \\ &+ \beta_{12} dgnp_t + \beta_{13} FAAB_t + \beta_{14} Public_t + \beta_{15} ruralare_t + \beta_{16} largcity_t \\ &+ \beta_{17} market_t + \beta_{18} marketw_t\end{aligned}$$

Second stage

$$\begin{aligned}EDEP_t^{BU} = \gamma_0 &+ \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 dcat_{t-1} + \gamma_4 dcl_{t-1} + \gamma_5 EDEP_t^{MA} + \gamma_6 (EDEP_t^{MA})^2 \\ &+ \gamma_7 S_t^{MA} + \gamma_8 (S_t^{MA})^2 + \gamma_9 I_t^{MA} + \gamma_{10} BU_{t-1} + \gamma_{11} ddmpa_{t-1} \\ &+ \gamma_{12} dgnp_t + \gamma_{13} FAAB_t + \gamma_{14} Public_t + \gamma_{15} ruralare_t + \gamma_{16} largcity_t \\ &+ \gamma_{17} market_t + \gamma_{18} marketw_t\end{aligned}$$

2.3.5 IBU

First stage

$$\begin{aligned}\eta_{IBU} = \beta_0 &+ \beta_1 ddmcash_{t-1} + \beta_2 dcash_{t-1} + \beta_3 EDEP_t^{MA} + \beta_4 (EDEP_t^{MA})^2 + \beta_5 S_t^{MA} + \beta_6 I_t^{MA} \\ &+ \beta_7 EDEP_t^{BU} + \beta_8 (EDEP_t^{BU})^2 + \beta_9 dca_{t-1} + \beta_{10} ddmpa_{t-1} + \beta_{11} (ddmpa_{t-1})^2 \\ &+ \beta_{12} dcl_{t-1} + \beta_{13} dgnp_t + \beta_{14} FAAB_t + \beta_{15} Public_t + \beta_{16} ruralare_t \\ &+ \beta_{17} largcity_t + \beta_{18} market_t + \beta_{19} marketw_t\end{aligned}$$

Second stage

$$\begin{aligned}I_t^{BU} = \gamma_0 &+ \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 dcat_{t-1} + \gamma_4 dcl_{t-1} + \gamma_5 EDEP_t^{MA} + \gamma_6 (EDEP_t^{MA})^2 \\ &+ \gamma_7 S_t^{MA} + \gamma_8 I_t^{MA} + \gamma_9 EDEP_t^{BU} + \gamma_{10} (EDEP_t^{BU})^2 + \gamma_{11} ddmpa_{t-1} \\ &+ \gamma_{12} dgnp_t + \gamma_{13} FAAB_t + \gamma_{14} Public_t + \gamma_{15} ruralare_t + \gamma_{16} largcity_t \\ &+ \gamma_{17} market_t + \gamma_{18} marketw_t\end{aligned}$$

2.3.6 dofa

First stage

$$\begin{aligned}\eta_{dofa,pos} = \beta_0 &+ \beta_1 ddmcash_{t-1} + \beta_2 dcash_{t-1} + \beta_3 ddmpa_{t-1} + \beta_4 (ddmpa_{t-1})^2 + \beta_5 (ddmpa_{t-1})^3 \\ &+ \beta_6 DI_t^{MA} + \beta_7 DI_t^{BU} + \beta_8 realr_t + \beta_9 FAAB_t + \beta_{10} Public_t \\ &+ \beta_{11} ruralare_t + \beta_{12} largcity_t + \beta_{13} market_t + \beta_{14} marketw_t\end{aligned}$$

$$
\begin{aligned}
\eta_{dofa,neg} = \delta_0 &+ \delta_1 ddmcash_{t-1} + \delta_2 dcash_{t-1} + \delta_3 ddmpa_{t-1} + \delta_4 (ddmpa_{t-1})^2 + \delta_5 DI_t^{MA} \\
&+ \delta_6 DI_t^{BU} + \delta_7 realr_t + \delta_8 FAAB_t + \delta_9 Public_t + \delta_{10} ruralare_t \\
&+ \delta_{11} largcity_t + \delta_{12} market_t + \delta_{13} marketw_t
\end{aligned}
$$

Second stage for positive level

$$
\begin{aligned}
dofa_t = \gamma_0 &+ \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 ddmpa_{t-1} + \gamma_4 (ddmpa_{t-1})^2 + \gamma_5 (ddmpa_{t-1})^3 \\
&+ \gamma_6 DI_t^{MA} + \gamma_7 DI_t^{BU} + \gamma_8 realr_t + \gamma_9 FAAB_t + \gamma_{10} Public_t \\
&+ \gamma_{11} ruralare_t + \gamma_{12} largcity_t + \gamma_{13} market_t + \gamma_{14} marketw_t
\end{aligned}
$$

Second stage for negative level

$$
\begin{aligned}
dofa_t = \lambda_0 &+ \lambda_1 ddmcash_{t-1} + \lambda_2 dcash_{t-1} + \lambda_3 ddmpa_{t-1} + \lambda_4 (ddmpa_{t-1})^2 + \lambda_5 DI_t^{MA} \\
&+ \lambda_6 DI_t^{BU} + \lambda_7 realr_t + \lambda_8 FAAB_t + \lambda_9 Public_t + \lambda_{10} ruralare_t \\
&+ \lambda_{11} largcity_t + \lambda_{12} market_t + \lambda_{13} marketw_t
\end{aligned}
$$

2.3.7 dca

$$
\begin{aligned}
dca_t = \gamma_0 &+ \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 EDEP_t^{MA} + \gamma_4 (EDEP_t^{MA})^2 + \gamma_5 S_t^{MA} + \gamma_6 I_t^{MA} \\
&+ \gamma_7 EDEP_t^{BU} + \gamma_8 (EDEP_t^{BU})^2 + \gamma_9 I_t^{BU} + \gamma_{10} (I_t^{BU})^2 + \gamma_{11} (I_t^{BU})^3 \\
&+ \gamma_{12} dcl_t + \gamma_{13} ddmpa_{t-1} + \gamma_{14} (ddmpa_{t-1})^2 + \gamma_{15} dgnp_t + \gamma_{16} FAAB_t \\
&+ \gamma_{17} Public_t + \gamma_{18} ruralare_t + \gamma_{19} largcity_t + \gamma_{20} market_t + \gamma_{21} marketw_t
\end{aligned}
$$

2.3.8 dll

First stage

$$
\begin{aligned}
\eta_{dll} = \beta_0 &+ \beta_1 ddmcash_{t-1} + \beta_2 dcash_{t-1} + \beta_3 ddmpa_{t-1} + \beta_4 (ddmpa_{t-1})^2 + \beta_5 (ddmpa_{t-1})^3 \\
&+ \beta_6 DI_t^{MA} + \beta_7 DI_t^{BU} + \beta_8 Ddofa_t + \beta_9 realr_t + \beta_{10} FAAB_t \\
&+ \beta_{11} Public_t + \beta_{12} ruralare_t + \beta_{13} largcity_t + \beta_{14} market_t + \beta_{15} marketw_t
\end{aligned}
$$

Second stage

$$
\begin{aligned}
dll_t = \gamma_0 &+ \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 ddmpa_{t-1} + \gamma_4 (ddmpa_{t-1})^2 + \gamma_5 (ddmpa_{t-1})^3 \\
&+ \gamma_6 DI_t^{MA} + \gamma_7 DI_t^{BU} + \gamma_8 Ddofa_t + \gamma_9 realr_t + \gamma_{10} FAAB_t \\
&+ \gamma_{11} Public_t + \gamma_{12} ruralare_t + \gamma_{13} largcity_t + \gamma_{14} market_t + \gamma_{15} marketw_t
\end{aligned}
$$

2.3.9 dcl

$$
\begin{aligned}
dcl_t = \gamma_0 &+ \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 EDEP_t^{MA} + \gamma_4 (EDEP_t^{MA})^2 + \gamma_5 S_t^{MA} + \gamma_6 I_t^{MA} \\
&+ \gamma_7 EDEP_t^{BU} + \gamma_8 (EDEP_t^{BU})^2 + \gamma_9 I_t^{BU} + \gamma_{10} (I_t^{BU})^2 + \gamma_{11} ddmpa_{t-1} \\
&+ \gamma_{12} (ddmpa_{t-1})^2 + \gamma_{13} (ddmpa_{t-1})^3 + \gamma_{14} dca_t + \gamma_{15} dgnp_t + \gamma_{16} FAAB_t \\
&+ \gamma_{17} Public_t + \gamma_{18} ruralare_t + \gamma_{19} largcity_t + \gamma_{20} market_t + \gamma_{21} marketw_t
\end{aligned}
$$

2.3.10 dsc

First stage

$$
\begin{aligned}
\eta_{dsc,pos} = \beta_0 &+ \beta_1 ddmcash_{t-1} + \beta_2 dcash_{t-1} + \beta_3 ddmpa_{t-1} + \beta_4 (ddmpa_{t-1})^2 + \beta_5 (ddmpa_{t-1})^3 \\
&+ \beta_6 DI_t^{MA} + \beta_7 DI_t^{BU} + \beta_8 Ddofa_t + \beta_9 Ddll_t + \beta_{10} realr_t \\
&+ \beta_{11} FAAB_t + \beta_{12} Public_t + \beta_{13} ruralare_t + \beta_{14} largcity_t + \beta_{15} market_t + \beta_{16} marketw_t
\end{aligned}
$$

$$
\begin{aligned}
\eta_{dsc,neg} = \delta_0 &+ \delta_1 ddmcash_{t-1} + \delta_2 dcash_{t-1} + \delta_3 ddmpa_{t-1} + \delta_4 (ddmpa_{t-1})^2 + \delta_5 (ddmpa_{t-1})^3 \\
&+ \delta_6 DI_t^{MA} + \delta_7 DI_t^{BU} + \delta_8 Ddofa_t + \delta_9 Ddll_t + \delta_{10} realr_t \\
&+ \delta_{11} FAAB_t + \delta_{12} Public_t + \delta_{13} ruralare_t + \delta_{14} largcity_t + \delta_{15} market_t + \delta_{16} marketw_t
\end{aligned}
$$

Second stage for positive level

$$\begin{aligned} dsc_t = {} & \gamma_0 + \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 ddmpa_{t-1} + \gamma_4 (ddmpa_{t-1})^2 + \gamma_5 DI_t^{MA} \\ & + \gamma_6 DI_t^{BU} + \gamma_7 Ddofa_t + \gamma_8 Ddll_t + \gamma_9 realr_t + \gamma_{10} FAAB_t \\ & + \gamma_{11} Public_t + \gamma_{12} ruralare_t + \gamma_{13} largcity_t + \gamma_{14} market_t + \gamma_{15} marketw_t \end{aligned}$$

Second stage for negative level

$$\begin{aligned} dsc_t = {} & \lambda_0 + \lambda_1 ddmcash_{t-1} + \lambda_2 dcash_{t-1} + \lambda_3 ddmpa_{t-1} + \lambda_4 DI_t^{MA} + \lambda_5 DI_t^{BU} \\ & + \lambda_6 Ddofa_t + \lambda_7 Ddll_t + \lambda_8 realr_t + \lambda_9 FAAB_t + \lambda_{10} Public_t \\ & + \lambda_{11} ruralare_t + \lambda_{12} largcity_t + \lambda_{13} market_t + \lambda_{14} marketw_t \end{aligned}$$

2.3.11 drr

First stage

$$\begin{aligned} \eta_{drr} = {} & \beta_0 + \beta_1 ddmcash_{t-1} + \beta_2 (ddmcash_{t-1})^2 + \beta_3 DI_t^{MA} + \beta_4 DI_t^{BU} + \beta_5 Ddofa_t \\ & + \beta_6 Ddll_t + \beta_7 Ddsc_t + \beta_8 realr_t + \beta_9 FAAB_t + \beta_{10} Public_t \\ & + \beta_{11} ruralare_t + \beta_{12} largcity_t + \beta_{13} market_t + \beta_{14} marketw_t \end{aligned}$$

Second stage

$$\begin{aligned} drr_t = {} & \gamma_0 + \gamma_1 ddmcash_{t-1} + \gamma_2 (ddmcash_{t-1})^2 + \gamma_3 DI_t^{MA} + \gamma_4 DI_t^{BU} + \gamma_5 Ddofa_t \\ & + \gamma_6 Ddll_t + \gamma_7 Ddsc_t + \gamma_8 realr_t + \gamma_9 FAAB_t + \gamma_{10} Public_t \\ & + \gamma_{11} ruralare_t + \gamma_{12} largcity_t + \gamma_{13} market_t + \gamma_{14} marketw_t \end{aligned}$$

2.3.12 OIBD

$$\begin{aligned} OIBD_t = {} & \gamma_0 + \gamma_1 CA_{t-1} + \gamma_2 CL_{t-1} + \gamma_3 MA_{t-1} + \gamma_4 I_t^{MA} + \gamma_5 S_t^{MA} \\ & + \gamma_6 EDEP_t^{MA} + \gamma_7 (EDEP_t^{MA})^2 + \gamma_8 BU_{t-1} + \gamma_9 I_t^{BU} + \gamma_{10} EDEP_t^{BU} \\ & + \gamma_{11} (EDEP_t^{BU})^2 + \gamma_{12} dca_t + \gamma_{13} (dca_t)^2 + \gamma_{14} ddmpa_{t-1} + \gamma_{15} (ddmpa_{t-1})^2 \\ & + \gamma_{16} (ddmpa_{t-1})^3 + \gamma_{17} dcash_{t-1} + \gamma_{18} (dcash_{t-1})^2 + \gamma_{19} dcl_t + \gamma_{20} dgnp_t \\ & + \gamma_{21} FAAB_t + \gamma_{22} Public_t + \gamma_{23} ruralare_t + \gamma_{24} largcity_t + \gamma_{25} market_t + \gamma_{26} marketw_t \end{aligned}$$

2.3.13 FI

First stage

$$\begin{aligned} \eta_{FI} = {} & \beta_0 + \beta_1 I_t^{BU} + \beta_2 EDEP_t^{MA} + \beta_3 (EDEP_t^{MA})^2 + \beta_4 S_t^{MA} + \beta_5 I_t^{MA} \\ & + \beta_6 (I_t^{MA})^2 + \beta_7 EDEP_t^{BU} + \beta_8 (EDEP_t^{BU})^2 + \beta_9 dca_t + \beta_{10} (dca_t)^2 \\ & + \beta_{11} dofa_t + \beta_{12} OFA_{t-1} + \beta_{13} CA_{t-1} + \beta_{14} MA_{t-1} + \beta_{15} BU_{t-1} \\ & + \beta_{16} realr_t + \beta_{17} FAAB_t + \beta_{18} Public_t + \beta_{19} ruralare_t + \beta_{20} largcity_t \\ & + \beta_{21} market_t + \beta_{22} marketw_t \end{aligned}$$

Second stage

$$\begin{aligned} FI_t = {} & \gamma_0 + \gamma_1 I_t^{BU} + \gamma_2 EDEP_t^{MA} + \gamma_3 (EDEP_t^{MA})^2 + \gamma_4 S_t^{MA} + \gamma_5 I_t^{MA} \\ & + \gamma_6 (I_t^{MA})^2 + \gamma_7 EDEP_t^{BU} + \gamma_8 (EDEP_t^{BU})^2 + \gamma_9 dca_t + \gamma_{10} (dca_t)^2 \\ & + \gamma_{11} dofa_t + \gamma_{12} OFA_{t-1} + \gamma_{13} CA_{t-1} + \gamma_{14} MA_{t-1} + \gamma_{15} BU_{t-1} \\ & + \gamma_{16} realr_t + \gamma_{17} FAAB_t + \gamma_{18} Public_t + \gamma_{19} ruralare_t + \gamma_{20} largcity_t \\ & + \gamma_{21} market_t + \gamma_{22} marketw_t \end{aligned}$$

2.3.14 FE

First stage

$$\begin{aligned}\eta_{FE} = \beta_0 &+ \beta_1 I_t^{BU} + \beta_2 EDEP_t^{MA} + \beta_3 S_t^{MA} + \beta_4 I_t^{MA} + \beta_5 EDEP_t^{BU}\\ &+ \beta_6 OFA_{t-1} + \beta_7 MA_{t-1} + \beta_8 BU_{t-1} + \beta_9 LL_{t-1} + \beta_{10} CA_{t-1}\\ &+ \beta_{11} CL_{t-1} + \beta_{12} dca_t + \beta_{13} dcl_t + \beta_{14} dofa_t + \beta_{15} dll_t\\ &+ \beta_{16} realr_t + \beta_{17} FAAB_t + \beta_{18} Public_t + \beta_{19} ruralare_t + \beta_{20} largcity_t\\ &+ \beta_{21} market_t + \beta_{22} marketw_t\end{aligned}$$

Second stage

$$\begin{aligned}FE_t = \gamma_0 &+ \gamma_1 I_t^{BU} + \gamma_2 EDEP_t^{MA} + \gamma_3 S_t^{MA} + \gamma_4 I_t^{MA} + \gamma_5 EDEP_t^{BU}\\ &+ \gamma_6 OFA_{t-1} + \gamma_7 MA_{t-1} + \gamma_8 BU_{t-1} + \gamma_9 LL_{t-1} + \gamma_{10} CA_{t-1}\\ &+ \gamma_{11} CL_{t-1} + \gamma_{12} dca_t + \gamma_{13} dcl_t + \gamma_{14} dofa_t + \gamma_{15} dll_t\\ &+ \gamma_{16} realr_t + \gamma_{17} FAAB_t + \gamma_{18} Public_t + \gamma_{19} ruralare_t + \gamma_{20} largcity_t\\ &+ \gamma_{21} market_t + \gamma_{22} marketw_t\end{aligned}$$

2.3.15 TDEPMA

$$\begin{aligned}TDEP_t^{MA*} &= \beta_0 + \beta_1 ddmcash_{t-1} + \beta_2 dcash_{t-1} + \beta_3 S_t^{MA} + \beta_4 EDEP_t^{MA} + \beta_5 (EDEP_t^{MA})^2\\ &\quad + \beta_6 I_t^{MA} + \beta_7 (I_t^{MA})^2 + \beta_8 ddmpa_{t-1} + \beta_9 (ddmpa_{t-1})^2 + \beta_{10} (ddmpa_{t-1})^3\\ &\quad + \beta_{11} realr_t + \beta_{12} FAAB_t + \beta_{13} Public_t + \beta_{14} ruralare_t + \beta_{15} largcity_t\\ &\quad + \beta_{16} market_t + \beta_{17} marketw_t\\ TDEP_t^{MA} &= \max(0, TDEP_t^{MA*})\end{aligned}$$

2.3.16 zpf

First stage

$$\begin{aligned}\eta_{ZPF} = \beta_0 &+ \beta_1 ddmcash_{t-1} + \beta_2 dcash_{t-1} + \beta_3 p_{t-1}^{allo} + \beta_4 ddmpa_{t-1} + \beta_5 (ddmpa_{t-1})^2\\ &+ \beta_6 (ddmpa_{t-1})^3 + \beta_7 DTDEP_t^{MA} + \beta_8 realr_t + \beta_9 FAAB_t + \beta_{10} Public_t\\ &+ \beta_{11} ruralare_t + \beta_{12} largcity_t + \beta_{13} market_t + \beta_{14} marketw_t\end{aligned}$$

Second stage

$$\begin{aligned}zpf_t = \gamma_0 &+ \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 p_{t-1}^{allo} + \gamma_4 ddmpa_{t-1} + \gamma_5 DTDEP_t^{MA}\\ &+ \gamma_6 realr_t + \gamma_7 FAAB_t + \gamma_8 Public_t + \gamma_9 ruralare_t + \gamma_{10} largcity_t\\ &+ \gamma_{11} market_t + \gamma_{12} marketw_t\end{aligned}$$

2.3.17 dour

First stage

$$\begin{aligned}\eta_1 = Intercept_1 + Intercept_2 &+ \beta_{1,1} ddmcash_{t-1} + \beta_{2,1} dcash_{t-1} + \beta_{3,1} ddmpa_{t-1} + \beta_{4,1} (ddmpa_{t-1})^2\\ &+ \beta_{5,1} (ddmpa_{t-1})^3 + \beta_{6,1} DTDEP_t^{MA} + \beta_{7,1} Dzpf_t + \beta_{8,1} realr_t\\ &+ \beta_{9,1} FAAB_t + \beta_{10,1} Public_t + \beta_{11,1} ruralare_t + \beta_{12,1} largcity_t\\ &+ \beta_{13,1} market_t + \beta_{14,1} marketw_t\end{aligned}$$

Second stage for positive

$$\begin{aligned}dour_t = \gamma_0 &+ \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 ddmpa_{t-1} + \gamma_4 (ddmpa_{t-1})^2 + \gamma_5 DTDEP_t^{MA}\\ &+ \gamma_6 Dzpf_t + \gamma_7 realr_t + \gamma_8 FAAB_t + \gamma_9 Public_t + \gamma_{10} ruralare_t\\ &+ \gamma_{11} largcity_t + \gamma_{12} market_t + \gamma_{13} marketw_t\end{aligned}$$

Second stage for negative

$$\begin{aligned} dour_t = \delta_0 &+ \delta_1 ddmcash_{t-1} + \delta_2 dcash_{t-1} + \delta_3 ddmpa_{t-1} + \delta_4 DTDEP_t^{MA} + \delta_5 Dzpf_t \\ &+ \delta_6 realr_t + \delta_7 FAAB_t + \delta_8 Public_t + \delta_9 ruralare_t + \delta_{10} largcity_t \\ &+ \delta_{11} market_t + \delta_{12} marketw_t \end{aligned}$$

2.3.18 GC

First stage

$$\begin{aligned} \eta_{GC,pos} = \beta_0 &+ \beta_1 OIBD_t + \beta_2 (OIBD_t)^2 + \beta_3 (OIBD_t)^3 + \beta_4 FI_t + \beta_5 FE_t \\ &+ \beta_6 TDEP_t^{MA} + \beta_7 (TDEP_t^{MA})^2 + \beta_8 EDEP_t^{BU} + \beta_9 (EDEP_t^{BU})^2 + \beta_{10} zpf_t \\ &+ \beta_{11} dour_t + \beta_{12} dgnp_t + \beta_{13} FAAB_t + \beta_{14} Public_t + \beta_{15} ruralare_t \\ &+ \beta_{16} largcity_t + \beta_{17} market_t + \beta_{18} marketw_t \end{aligned}$$

$$\begin{aligned} \eta_{GC,neg} = \delta_0 &+ \delta_1 OIBD_t + \delta_2 (OIBD_t)^2 + \delta_3 FI_t + \delta_4 FE_t + \delta_5 TDEP_t^{MA} \\ &+ \delta_6 (TDEP_t^{MA})^2 + \delta_7 EDEP_t^{BU} + \delta_8 (EDEP_t^{BU})^2 + \delta_9 zpf_t + \delta_{10} dour_t \\ &+ \delta_{11} dgnp_t + \delta_{12} FAAB_t + \delta_{13} Public_t + \delta_{14} ruralare_t + \delta_{15} largcity_t \\ &+ \delta_{16} market_t + \delta_{17} marketw_t \end{aligned}$$

Second stage for positive

$$\begin{aligned} GC_t = \gamma_0 &+ \gamma_1 OIBD_t + \gamma_2 (OIBD_t)^2 + \gamma_3 (OIBD_t)^3 + \gamma_4 FI_t + \gamma_5 FE_t \\ &+ \gamma_6 TDEP_t^{MA} + \gamma_7 (TDEP_t^{MA})^2 + \gamma_8 EDEP_t^{BU} + \gamma_9 (EDEP_t^{BU})^2 + \gamma_{10} zpf_t \\ &+ \gamma_{11} dour_t + \gamma_{12} dgnp_t + \gamma_{13} FAAB_t + \gamma_{14} Public_t + \gamma_{15} ruralare_t \\ &+ \gamma_{16} largcity_t + \gamma_{17} market_t + \gamma_{18} marketw_t \end{aligned}$$

Second stage for negative

$$\begin{aligned} GC_t = \lambda_0 &+ \lambda_1 OIBD_t + \lambda_2 (OIBD_t)^2 + \lambda_3 FI_t + \lambda_4 FE_t + \lambda_5 TDEP_t^{MA} \\ &+ \lambda_6 (TDEP_t^{MA})^2 + \lambda_7 EDEP_t^{BU} + \lambda_8 (EDEP_t^{BU})^2 + \lambda_9 zpf_t + \lambda_{10} dour_t \\ &+ \lambda_{11} dgnp_t + \lambda_{12} FAAB_t + \lambda_{13} Public_t + \lambda_{14} ruralare_t + \lambda_{15} largcity_t \\ &+ \lambda_{16} market_t + \lambda_{17} marketw_t \end{aligned}$$

2.3.19 OA

First stage

$$\begin{aligned} \eta_1 = Intercept_1 + Intercept_2 + \beta_{0,1} &+ \beta_{1,1} dour_t + \beta_{2,1} GC_t + \beta_{3,1} DTDEP_t^{MA} + \beta_{4,1} Dzpf_t \\ &+ \beta_{5,1} realr_t + \beta_{6,1} FAAB_t + \beta_{7,1} Public_t + \beta_{8,1} ruralare_t \\ &+ \beta_{9,1} largcity_t + \beta_{10,1} market_t + \beta_{11,1} marketw_t \end{aligned}$$

Second stage for positive

$$\begin{aligned} OA_t = \gamma_0 &+ \gamma_1 dour_t + \gamma_2 GC_t + \gamma_3 DTDEP_t^{MA} + \gamma_4 Dzpf_t + \gamma_5 realr_t \\ &+ \gamma_6 FAAB_t + \gamma_7 Public_t + \gamma_8 ruralare_t + \gamma_9 largcity_t + \gamma_{10} market_t + \gamma_{11} marketw_t \end{aligned}$$

Second stage for negative

$$\begin{aligned} OA_t = \delta_0 &+ \delta_1 dour_t + \delta_2 GC_t + \delta_3 DTDEP_t^{MA} + \delta_4 Dzpf_t + \delta_5 realr_t \\ &+ \delta_6 FAAB_t + \delta_7 Public_t + \delta_8 ruralare_t + \delta_9 largcity_t + \delta_{10} market_t + \delta_{11} marketw_t \end{aligned}$$

2.3.20 TL

First stage

$$
\begin{aligned}
\eta_{TL} = \beta_0 &+ \beta_1 OIBD_t + \beta_2 (OIBD_t)^2 + \beta_3 FI_t + \beta_4 (FI_t)^2 + \beta_5 FE_t \\
&+ \beta_6 (FE_t)^2 + \beta_7 TDEP_t^{MA} + \beta_8 (TDEP_t^{MA})^2 + \beta_9 EDEP_t^{BU} + \beta_{10} (EDEP_t^{BU})^2 \\
&+ \beta_{11} dour_t + \beta_{12} (dour_t)^2 + \beta_{13} zpf_t + \beta_{14} p_{t-1}^{allo} + \beta_{15} dgnp_t \\
&+ \beta_{16} FAAB_t + \beta_{17} Public_t + \beta_{18} ruralare_t + \beta_{19} largcity_t + \beta_{20} market_t + \beta_{21} marketw_t
\end{aligned}
$$

Second stage

$$
\begin{aligned}
TL_t = \gamma_0 &+ \gamma_1 OIBD_t + \gamma_2 (OIBD_t)^2 + \gamma_3 FI_t + \gamma_4 (FI_t)^2 + \gamma_5 FE_t \\
&+ \gamma_6 (FE_t)^2 + \gamma_7 TDEP_t^{MA} + \gamma_8 (TDEP_t^{MA})^2 + \gamma_9 EDEP_t^{BU} + \gamma_{10} (EDEP_t^{BU})^2 \\
&+ \gamma_{11} dour_t + \gamma_{12} (dour_t)^2 + \gamma_{13} zpf_t + \gamma_{14} p_{t-1}^{allo} + \gamma_{15} dgnp_t \\
&+ \gamma_{16} FAAB_t + \gamma_{17} Public_t + \gamma_{18} ruralare_t + \gamma_{19} largcity_t + \gamma_{20} market_t + \gamma_{21} marketw_t
\end{aligned}
$$

2.3.21 OTA

First stage

$$
\begin{aligned}
\eta_1 = Intercept_1 + Intercept_2 &+ \beta_{1,1} p_{t-1}^{allo} + \beta_{2,1} zpf_t + \beta_{3,1} TDEP_t^{MA} + \beta_{4,1} (TDEP_t^{MA})^2 \\
&+ \beta_{5,1} OIBD_t + \beta_{6,1} (OIBD_t)^2 + \beta_{7,1} EDEP_t^{BU} + \beta_{8,1} (EDEP_t^{BU})^2 \\
&+ \beta_{9,1} dour_t + \beta_{10,1} TL_t + \beta_{11,1} FI_t + \beta_{12,1} FE_t \\
&+ \beta_{13,1} dgnp_t + \beta_{14,1} FAAB_t + \beta_{15,1} Public_t + \beta_{16,1} ruralare_t \\
&+ \beta_{17,1} largcity_t + \beta_{18,1} market_t + \beta_{19,1} marketw_t
\end{aligned}
$$

Second stage for positive

$$
\begin{aligned}
OTA_t = \gamma_0 &+ \gamma_1 p_{t-1}^{allo} + \gamma_2 zpf_t + \gamma_3 TDEP_t^{MA} + \gamma_4 (TDEP_t^{MA})^2 + \gamma_5 OIBD_t \\
&+ \gamma_6 (OIBD_t)^2 + \gamma_7 EDEP_t^{BU} + \gamma_8 (EDEP_t^{BU})^2 + \gamma_9 dour_t + \gamma_{10} TL_t \\
&+ \gamma_{11} FI_t + \gamma_{12} FE_t + \gamma_{13} dgnp_t + \gamma_{14} FAAB_t + \gamma_{15} Public_t \\
&+ \gamma_{16} ruralare_t + \gamma_{17} largcity_t + \gamma_{18} market_t + \gamma_{19} marketw_t
\end{aligned}
$$

Second stage for negative

$$
\begin{aligned}
OTA_t = \delta_0 &+ \delta_1 p_{t-1}^{allo} + \delta_2 zpf_t + \delta_3 TDEP_t^{MA} + \delta_4 OIBD_t + \delta_5 (OIBD_t)^2 \\
&+ \delta_6 EDEP_t^{BU} + \delta_7 (EDEP_t^{BU})^2 + \delta_8 dour_t + \delta_9 TL_t + \delta_{10} FI_t \\
&+ \delta_{11} FE_t + \delta_{12} dgnp_t + \delta_{13} FAAB_t + \delta_{14} Public_t + \delta_{15} ruralare_t \\
&+ \delta_{16} largcity_t + \delta_{17} market_t + \delta_{18} marketw_t
\end{aligned}
$$

2.3.22 TDEPBU

$$
\begin{aligned}
TDEP_t^{BU*} = \beta_0 &+ \beta_1 ddmcash_{t-1} + \beta_2 dcash_{t-1} + \beta_3 EDEP_t^{MA} + \beta_4 (EDEP_t^{MA})^2 + \beta_5 S_t^{MA} \\
&+ \beta_6 I_t^{MA} + \beta_7 BU_{t-1} + \beta_8 (BU_{t-1})^2 + \beta_9 dca_t + \beta_{10} (dca_t)^2 \\
&+ \beta_{11} dcl_t + \beta_{12} ddmpa_{t-1} + \beta_{13} (ddmpa_{t-1})^2 + \beta_{14} (ddmpa_{t-1})^3 + \beta_{15} dgnp_t \\
&+ \beta_{16} FAAB_t + \beta_{17} Public_t + \beta_{18} ruralare_t + \beta_{19} largcity_t + \beta_{20} market_t + \beta_{21} marketw_t \\
TDEP_t^{BU} &= \max(0, TDEP_t^{BU*})
\end{aligned}
$$

2.3.23 pallo

$$
\begin{aligned}
p_t^{allo} = \gamma_0 &+ \gamma_1 ddmcash_{t-1} + \gamma_2 dcash_{t-1} + \gamma_3 zpf_t + \gamma_4 dmpa_{t-1} + \gamma_5 MPA_t \\
&+ \gamma_6 realr_t + \gamma_7 FAAB_t + \gamma_8 Public_t + \gamma_9 ruralare_t + \gamma_{10} largcity_t \\
&+ \gamma_{11} market_{t-1} + \gamma_{12} marketw_{t-1}
\end{aligned}
$$

2.3.24 ROT

First stage

$$
\begin{aligned}
\eta_{ROT} = {} & \beta_0 + \beta_1 p_t^{allo} + \beta_2 zpf_t + \beta_3 TDEP_t^{MA} + \beta_4 (TDEP_t^{MA})^2 + \beta_5 OIBD_t \\
& + \beta_6 (OIBD_t)^2 + \beta_7 EDEP_t^{BU} + \beta_8 (EDEP_t^{BU})^2 + \beta_9 OTA_t + \beta_{10} (OTA_t)^2 \\
& + \beta_{11} TDEP_t^{BU} + \beta_{12} (TDEP_t^{BU})^2 + \beta_{13} dour_t + \beta_{14} TL_t + \beta_{15} FI_t \\
& + \beta_{16} FE_t + \beta_{17} dgnp_t + \beta_{18} FAAB_t + \beta_{19} Public_t + \beta_{20} ruralare_t \\
& + \beta_{21} largcity_t + \beta_{22} market_t + \beta_{23} marketw_t
\end{aligned}
$$

Second stage

$$
\begin{aligned}
ROT_t = {} & \gamma_0 + \gamma_1 p_t^{allo} + \gamma_2 zpf_t + \gamma_3 OIBD_t + \gamma_4 EDEP_t^{BU} + \gamma_5 TDEP_t^{BU} \\
& + \gamma_6 dour_t + \gamma_7 TL_t + \gamma_8 FI_t + \gamma_9 FE_t + \gamma_{10} dgnp_t \\
& + \gamma_{11} FAAB_t + \gamma_{12} Public_t + \gamma_{13} ruralare_t + \gamma_{14} largcity_t + \gamma_{15} market_t + \gamma_{16} marketw_t
\end{aligned}
$$

3 Synthetic data

The article simulates a single path of the dynamic system. To obtain the synthetic data, we need to assume noises for behaviors.

3.1 Preliminary

To start, we need the following:

1. Behavior parameters
2. Noise terms and co-variance matrix at $t = 0, 1, 2, 3$.
3. Mean values and covariance of flow variables and stock variables at $t = 0$.

3.1.1 Parameters

All estimated parameters are listed in Table 1-24

3.1.2 Noises

The co-variance matrix of noises are assumed to be diagonal.

3.1.3 Summaries of methods, noises distribution

Dependent	Method	States	First state	Noise of first state	Second state	Noise of second state
EDEPMA (T1)	LLG	Two	Logistic	clog-log	Huber-Schweppes	Heavy tail
SMA (T2)	MUNO	Three	Multinomial	clog-log	Huber-Schweppes	Heavy tail
IMA (T3)	Tobit	One	Tobit 1	Logistic	N/A	N/A
EDEPBU (T4)	LLG	Two	Logistic	clog-log	Huber-Schweppes	Heavy tail
IBU (T5)	LLG	Two	Logistic	clog-log	Huber-Schweppes	Heavy tail
dofa (T6)	LSG	Three	Logistic	clog-log	Huber-Schweppes	Heavy tail
dca (T7)	HS	One	Huber-Schweppes	Heavy tail	N/A	N/A
dll (T8)	LLN	Two	Logistic	Cum. Logistic	Huber-Schweppes	Heavy tail
dcl (T9)	HS	One	Huber-Schweppes	Heavy tail	N/A	N/A
dsc (T10)	LSG	Three	Logistic	clog-log	Huber-Schweppes	Heavy tail
drr (T11)	LLG	Two	Logistic	clog-log	Huber-Schweppes	Heavy tail
OIBD (T12)	HS	One	Huber-Schweppes	Heavy tail	N/A	N/A
FI (T13)	LLG	Two	Logistic	clog-log	Huber-Schweppes	Heavy tail
FE (T14)	LLG	Two	Logistic	clog-log	Huber-Schweppes	Heavy tail
TDEPMA (T15)	Tobit	One	Tobit 1	Logistic	N/A	N/A
ZPF (T16)	LLG	Two	Logistic	clog-log	Huber-Schweppes	Heavy tail
dour (T17)	MUNO	Three	Multinomial	clog-log	Huber-Schweppes	Heavy tail
GC (T18)	LSG	Three	Logistic	clog-log	Huber-Schweppes	Heavy tail
OA (T19)	MUNO	Three	Multinomial	clog-log	Huber-Schweppes	Heavy tail
TL (T20)	LLN	Two	Logistic	Cum. Logistic	Huber-Schweppes	Heavy tail
OTA (T21)	MUNO	Three	Multinomial	clog-log	Huber-Schweppes	Heavy tail
TDEPBU (T22)	Tobit	One	Tobit 1	Logistic	N/A	N/A
Pallo (T23)	HS	One	Huber-Schweppes	Heavy tail	N/A	N/A
ROT (T24)	LLG	Two	Logistic	clog-log	Huber-Schweppes	Heavy tail

3.2 Synthetic data generation

3.2.1 Variables at $t = 0$ method 1

Detailed process is as following:

1. Generate N companies.

2. Assign exogenous features to each company, including $Largecity_t$, $Rural_t$.

3. For simplicity, macro exogenous variables $dgnp, realr$ are set to 0. Number of all closed firms are 0 so that $FAAB = 0$ for all observations.

4. There is only one sector so we set $market_t = 0$ as well, while $marketw_t$ are calculated using the way in the paper.

5. Follow Velez-Pareja (09)'s method to start a company.

3.2.2 Velez-Pareja (09)'s application

Financial statements evolves recursively. Start a company at year $t-1$(Starting from zero). Define BS_{t-1} and CB_{t-1}, and calculate Investment at $t-1$, then IS_t, then BS_t, then Investment in t.

It shares the same spirit as Shahnzarian (04). The key is how to construct a company at year $t-1$ while holding accounting identity:

- Initial Equity Invested ($URE_{t-1}+SC_{t-1}+RR_{t-1}$), which are unrestricted equity, shared capital and restricted reserves. Also need to assume minimum current assets (CA_{t-1})

- Fixed asset first

$$(MA_{t-1} = I^{MA}_{t-1}, BU_{t-1} = I^{BU}_{t-1}, OFA_{t-1} = dofa_{t-1})$$

- Financing for fixed assets,

$$CA_{t-1} = dca_{t-1} = CL_{t-1} = dcl_{t-1}$$

,

$$dll_{t-1} = LL_{t-1} = (MA_{t-1} + BU_{t-1} + OFA_{t-1}) - (URE_{t-1} + SC_{t-1} + RR_{t-1})$$

- Accumulative Supplementary Depreciation, Other untaxed reserves ($ASD_{t-1} = 0, OUR_{t-1} = 0$). So that

$$CMA_{t-1} = MA_{t-1}$$

- Other necessary flow variables at $t-1$:

$$EDEP^{BU}_{t-1} = TDEP^{BU}_{t-1} = p^{allo}_{t-1} = dcashfl_{t-1} = ddmpa_{t-1} = ddmcash_{t-1} = 0$$

- With information above, we can decide all flow variables at t.

- Then, we can start to simulate. After long enough simulation, we take the last several years as our synthetic data.

- Detailed assumption ratios:

$$URE : SC : RR = 10 : 3 : 3$$
$$CA : URE = 1 : 1$$
$$MA : BU : OFA = 1 : 1 : 6$$
$$MA : URE = 1 : 1.5$$

- Start with a given level of $dca = CA \sim Uniform[0.2, 0.6] \times 1million$.

3.2.3 Variables at $t = 1, 2, 3$

After having all synthetic data at $t = 0$, we can recursively update the variables. When perform updates, we use different methods for each variable.

For HS

1. Generate error $\tilde{\epsilon}_t$ using Gaussian mixture or pearson

2. Compute

$$\tilde{y}_i = X_i'\hat{\beta} + \tilde{\epsilon}_t$$

For Tobit model:

1. Generate error $\tilde{\epsilon}_t = scale \times ln\left(\frac{U}{1-U}\right)$ where $U \sim Uniform(0,1)$

2. Compute latent variable

$$\tilde{y}_i^* = X_i'\hat{\beta} + \tilde{\epsilon}_t$$

3. Censor variable to obtain the final prediction

$$\tilde{y}_i = max\,(0, \tilde{y}_i^*)$$

For logistic with cloglog model (LLG):
1. Generate error $\tilde{\epsilon}_t$ using Gaussian mixture or pearson
2. Check if a firm will make decision:

$$\hat{\eta} = X_i'\hat{\beta}_{1st}$$
$$\hat{P} = 1 - e^{-e^{\hat{\eta}}}$$
$$U \sim Uniform(0,1)$$

- If $\hat{P} > U$, then make decision
- If $\hat{P} < U$, then not make decision, and set $= 0$
3. Generate data for firms that make positive decision

$$\tilde{y} = X'\hat{\beta}_{2nd} + \epsilon_t$$

For logistic with cumlulative logistic model (LLN):
1. Generate error $\tilde{\epsilon}_t$ using Gaussian mixture or pearson
2. Check if a firm will make decision:

$$\hat{\eta} = X_i'\hat{\beta}_{1st}$$
$$\hat{P} = \frac{exp\,(\hat{\eta})}{1 + exp\,(\hat{\eta})}$$
$$U \sim Uniform(0,1)$$

- If $\hat{P} > U$, then make decision
- If $\hat{P} < U$, then not make decision, and set $= 0$
3. Generate data for firms that make positive decision

$$\tilde{y} = X_i'\hat{\beta}_{2nd} + \epsilon_t$$

For multinomial logit with cloglog model (MUNO):
1. Generate error $\tilde{\epsilon}_t$ using Gaussian mixture or pearson
2. Create the thresholds for firms and decide firms' action

$$\hat{\eta_1} = Intercept_1 + X'\hat{\beta}_{1st}$$
$$\hat{\eta_2} = Intercept_2 + X'\hat{\beta}_{1st}$$
$$\hat{P_1} = 1 - e^{-e^{\hat{\eta_1}}}$$
$$\hat{P_2} = 1 - e^{-e^{\hat{\eta_2}}}$$
$$U \sim Uniform(0,1)$$

where $Intercept_1$ and $Intercept_2$ come from the same estimation result table.
- If $U < \hat{P_1}$, then take negative action
- If $\hat{P_1} < U < \hat{P_2}$, then does not take any action, and set $= 0$
- If $\hat{P_2} < U$, then take positive action
3. Generate data for firms that make either positive or negative decision

$$\tilde{y}_i = X_i'\hat{\beta}_{2nd} + \epsilon_{it}$$

where $i = positive, negative$

For logistic with cloglog model (LSG):
1. Generate error $\tilde{\epsilon}_t$ using Gaussian mixture or pearson

2. Create the thresholds for firms and decide firms' action

$$
\begin{aligned}
\hat{\eta_1} &= Intercept_{pos} + X'_{pos}\hat{\beta}_{1st,pos} \\
\hat{\eta_2} &= Intercept_{neg} + X'_{neg}\hat{\beta}_{1st,neg} \\
\hat{P}_1 &= 1 - e^{-e^{\hat{\eta_1}}} \\
\hat{P}_2 &= 1 - e^{-e^{\hat{\eta_2}}} \\
U &\sim Uniform(0,1)
\end{aligned}
$$

where $Intercept_{pos}$ and X_{pos} come from one estimation result table and $Intercept_{neg}$ and X_{neg} come from the other estimation result table.

- If $U < \hat{P}_1$, then take positive action
- If $\hat{P}_1 < U < \hat{P}_2$, then does not take any action, and set $= 0$
- If $\hat{P}_2 < U$, then take negative action

3. Generate data for firms that make either positive or negative decision

$$\tilde{y_i} = X'_i\hat{\beta}_{2nd,i} + \epsilon_{it}$$

where $i = positive, negative$

3.3 Noise simulation

Suppose we have four moments: Mean, Variance, Skewness and Kurtosis. We want to find a distribution to fit these four moments. Two distribution are suggested: Gaussian mixture and JohnsonSU.

3.3.1 JohnsonSU

Random variable X follows JohnsonSU distribution if it follows

$$Z = \gamma + \delta sinh^{-1}\left(\frac{x-\xi}{\lambda}\right)$$

where $Z \sim N(0,1)$. It has 4 parameters: ξ for location, λ for scale, γ and δ for shape,

Parameter	
ξ	location
λ	scale
γ	shape
δ	shape

We have four inputs, $\mu(x)$, $\sigma(x)$, $Skew$, $Kurt$. The algorithm is as the following:

1. Define a variable called ω, whose value can be searched from an given interval. Define $\beta_2 = Kurt$,$\beta_1 = Skew^2$.

2. Find a lower boundary ω_1, which is a positive root of $\omega^4 + 2\omega^3 + 3\omega^2 - 3 = \beta_2$. The solution is

$$
\begin{aligned}
D &= (3+\beta_2)\left(16\beta_2^2 + 87\beta_2 + 171\right)/27 \\
d &= -1 + \sqrt[3]{7 + 2\beta_2 + 2\sqrt{D}} - \sqrt[3]{2\sqrt{D} - 7 - 2\beta_2} \\
\omega_1 &= \frac{1}{2} \times \left(-1 + \sqrt{d} + \sqrt{\frac{4}{\sqrt{d}} - d - 3}\right)
\end{aligned}
$$

3. Define $f(\omega)$ as

$$
\begin{aligned}
m(\omega) &= -2 + \sqrt{4 + 2\left[\omega^2 - \frac{\beta_2 + 3}{\omega^2 + 2\omega + 3}\right]} \\
f(\omega) &= (\omega - 1 - m(\omega))\left(\omega + 2 + \frac{1}{2}m(\omega)\right)^2
\end{aligned}
$$

4. If $(\omega_1 - 1)(\omega_1 + 2)^2 \leq \beta_1$, there is no solution, otherwise go to next step.

5. $\omega_2 = \sqrt{-1+\sqrt{2\left(\beta_2-1\right)}}$.
6. Use a root finding procedure on the interval $(\omega_1, \omega_2]$ to find the desired ω^* that satisfies:

$$f(\omega^*) = \beta_1$$

7. With ω^*, we can recover the parameters according to the following:

$$
\begin{aligned}
m &= m(\omega^*) \\
\Omega &= -sgn\left(Skew\right) sinh^{-1}\sqrt{\frac{\omega^*-1}{2\omega^*}\left(\frac{\omega^*-1}{m}-1\right)} \\
\delta &= 1/\sqrt{ln\omega^*} \\
\gamma &= \Omega/\sqrt{ln\omega^*} \\
\lambda &= \frac{\sigma\left(x\right)}{\omega^*-1}\sqrt{\frac{2m}{\omega^*+1}} \\
\xi &= \mu(x) - sgn\left(Skew\right)\frac{\sigma(x)}{\omega^*-1}\sqrt{\omega^*-1-m}
\end{aligned}
$$

where

$$sgn(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

3.3.2 Gaussian mixture

The JohnsonSU is the quickest way, but it only applies to cases with $\left(\omega_1-1\right)\left(\omega_1+2\right)^2 \leq \beta_1$. For the rest of the cases, we fit the following distribution

$$p\left(x|\theta\right) = \sum_{i=1}^{K} w_i N\left(x|\mu_i, \sigma_i^2\right)$$

where $\theta = (w_1, ...w_k, \mu_1, ..., \mu_k, \sigma_1, ...\sigma_k)$. We minimize the loss:

$$\left\{\begin{array}{c} logit\left(w_1\right), ..., logit\left(w_{K-1}\right) \\ \mu_1, ..., \mu_L \\ \sigma_1^2, ..., \sigma_K^2 \end{array}\right\} = argmin\left[\left(\frac{FittedSkew - TargetSkew}{FittedSkew}\right)^2 + \left(\frac{FittedKurt - TargetKurt}{FittedKurt}\right)^2\right]$$

and obtain $g(x)$.

Then, we do a linear transformation such that

$$y = ax + b$$

where

$$
\begin{aligned}
a &= \frac{\sigma_{target}}{\sigma_x} \\
b &= \mu_{target} - a\mu_x
\end{aligned}
$$

to make sure we reach our desired mean and variance. The final distribution to be used is

$$p_{final}\left(x|\theta\right) = a \times \sum_{i=1}^{K} w_i N\left(x|\mu_i, \sigma_i^2\right) + b$$

3.3.3 Gaussian mixture linear transformation proof

Note that we only need to fit Skewness and Kurtosis, and do a linear transform of the generated desired results without changing the Skewness and Kurtosis. Suppose Mean and Std of X is μ_x,σ_x, $y = ax + b$, $a > 0$

Proof:

Step 1: Obviously, with $\mu_y = a\mu_x + b$, $\sigma_y^2 = a^2\sigma_x^2$, $\sigma_y = |a|\sigma_x$, we know the target mean and variance are matched.

$$E(y) = E\left[ax + b\right] = a\mu_x + (\mu_{target} - a\mu_x) = \mu_{target}$$

$$Var\left(y\right) = Var\left(ax\right) = a^2 Var\left(x\right) = \frac{\sigma_{target}^2}{\sigma_x^2}\sigma_x^2 = \sigma_{target}^2$$

Step 2: For $n \geq 2$, higher moments are

$$\mu_{n,y} = E\left[(y - \mu_y)^n\right] = E\left[(ax + b - (a\mu_x + b))^n\right] = E\left[(a\left(x - \mu_x\right))^n\right] = a^n E\left[(x - \mu_x)^n\right] = a^n \mu_{n,x}$$

Step 3: Skew and Kurt

$$skew(y) = \frac{\mu_{3,y}}{\sigma_y^3} = \frac{a^3\mu_{3,x}}{|a|^3\sigma_x^3} = sgn\left(a\right) skew(x)$$

$$kurt(y) = \frac{\mu_{4,y}}{\sigma_y^4} = \frac{a^4\mu_{4,x}}{|a|^4\sigma_x^4} = kurt(x)$$

Step 4: Since $a = \sigma_{target}/\sigma_x > 0$, the results hold.

Step 5: y is still a random variable and follows a valid distribution. The CDF of y is

$$F_y(y) = P\left(y \leq y_0\right) = P\left(ax + b \leq y_0\right) = P\left(x \leq \frac{y_0 - b}{a}\right) = F_x\left(\frac{y_0 - b}{a}\right)$$

with pdf

$$f_y(y) = \frac{1}{a} f_x\left(\frac{y - b}{a}\right)$$

Let $z = \frac{y-b}{a}$, then $y = az + b$, $dy = adz$

$$\int_{-\infty}^{\infty} \frac{1}{a} f_x\left(\frac{y - b}{a}\right) dy = \int_{-\infty}^{\infty} \frac{1}{a} f_x\left(\frac{y - b}{a}\right) adz = \int_{-\infty}^{\infty} f_x\left(z\right) dz = 1$$

3.4 Moments

$$mean = E\left(X\right) = \mu$$

$$Var = E\left(X - \mu\right)^2 = \sigma^2$$

$$Skew = \frac{E\left(X - \mu\right)^3}{\sigma^3}$$

$$Kurt = \frac{E\left(X - \mu\right)^4}{\sigma^4}$$