Macroeconomic Model of Credit Losses for Two Loan Portfolios

Petr Gapko[1], Martin Šmíd[1]

Abstract

We propose a dynamic macroeconomic model of credit risk for multiple portfolios with two factors for each portfolio. We follow the common approach that the credit risk on a loan portfolio can be decomposed into a probability of default and a loss given default and assume that both are driven by two underlying factors: one common for all borrowers in the portfolio and one individual for each single borrower. Our model additionally to the current research estimates the interconnectedness of the portfolios through the risk factors and enhances the current research by introducing dynamics and incorporating the external (macroeconomic) influence. We estimate the model on a set of two large real estate loan portfolios (one residential and one commercial) and show how the portfolios are interconnected and how the credit risk is influenced by macroeconomic environment.

Keywords: credit risk, mortgage, loan portfolio, dynamic model, estimation, interconnectedness, cointegration

JEL Classification: G32

[1] Econometric Department, Institute of Information Theory and Automation, Czech Academy of Sciences

1. Introduction

At the end of the last decade, when the financial crisis fully hit the US economy, losses from real estate loans in the US increased ten times, compared with the period of economic growth ending in 2007. The aim of our paper is to investigate the risk factors, which drove the increase of delinquencies and charge-offs of residential and commercial real estate loans and estimate the interconnectedness of the residential and commercial mortgage portfolios.

The recent research clearly proved that there is an obvious relationship between the state of the economy and the credit risk. Hamerle et al. (Hamerle, Dartsch, Jobst, & Plank, 2011) showed on a bond portfolio the necessity of taking into account changes in macroeconomic environment. Similarly, (Sommar & Shahnazarian, 2009) used the vector error correction model to estimate the dependency of expected default frequency of a portfolio of nonfinancial listed companies on several macroeconomic factors, from which they found the most influencing the interest rate. The mentioned results are in line with the findings of Pesaran et al. (Pesaran, Schuermann, Treutler, & Weiner, 2003) or Virolainen (Virolainen, 2004).

To this end, we propose a credit risk model based on Merton's assumption that credit risk is driven by underlying risk factors (Merton, 1974). We derive a Merton-Vasicek type of the loss distribution (Vasicek, 1987) with several extensions, of which the most important is the inclusion of macroeconomic environment. Our model converts default (or delinquency) rate and a real experienced percentage loss (or loss given default) on a portfolio into underlying factors. Our proposed methodology is conceptually similar to the approaches of Frye (Frye, 2000), Pykhtin (Pykhtin, 2003), Jimenez & Mencia (Jimenez & Mencia, 2009) or Witzany (Witzany, 2011) with the difference that in our approach, we study the interconnectedness of all factors and take into account the relevance of external (macroeconomic) conditions. On the other hand the studies examining the relationship between the credit risk and the macroeconomic environment, e.g. Pesaran (Pesaran, Schuermann, Treutler, & Weiner, 2003), do not decompose the credit risk into a multi-factor matrix.

Our approach brings a further extension of the abovementioned frameworks in three ways. First, we introduce dynamics in the underlying risk factors. Second, we use the cointegration analysis to find a relationship between common factors and the external (macroeconomic) environment. Also the approach to the estimation of the interconnectedness of multiple portfolios is an enhancement.

We estimate the model on a dataset of US nationwide residential and commercial real estate loan portfolios 30+ delinquencies (loans more than 30 days past due) and charge-off (net charge-offs of loans from books) rates. The cointegration analysis of underlying risk factors and macroeconomic variables such as GDP, unemployment, HPI, personal income, etc… clearly shows that the risk performance of a loan portfolio is linked to the macroeconomic environment. Additionally, our analysis shows that there exists a complex interconnectedness between loan portfolios, commercial and residential mortgages in our case.

The paper is organized as follows. In the following section we provide a description of the model methodology. In Section 3 we describe the data, the empirical analysis and our results. Finally, Section 4 concludes.

2. The Model

Similarly to Vasicek (Vasicek, 1987), we assume that the default of a loan happens when

$$A < B$$

where A is the value the debtor's (hypothetical) assets and B is the value of his debts. The recovery rate is, in line with Pykhtin (Pykhtin, 2003) computed as

$$R = \frac{\min(P,p)}{p} = \min(p^{-1}P, 1)$$

where p is the outstanding principle of the loan and P is the price of the collateral.

As it is usual, we assume that

$$a = \exp\{X + Z\}, \quad P = \exp\{H + E\}$$

where X, H are the common factors, and Z, E are mutually independent normally distributed individual factors, specific for each loan. For simplicity, we assume that B is common for all loans in the portfolio.

If the loan portfolio is large and homogeneous, then the default rate (sometimes imprecisely referenced as probability of default - PD) of a large homogeneous loan portfolio may be approximated as

$$Q = \frac{number\ of\ defaults}{number\ of\ loans} \doteq \mathbb{P}[A < B|Y] = \varphi(-Y), \qquad Y = \frac{X - \log B}{\rho}$$

where φ is a standard normal c.d.f. and ρ is the standard deviation of Z. The loss given default (LGD) - another quantity of usual interest - comes out as

$$G = \frac{total\ loss\ of\ porfolio}{number\ of\ defaults} \doteq \mathbb{E}[1 - R|I] = 1 - \mathbb{E}[\min(\exp\{I\ +\ E\}, 1)\ =\ \eta(I;\sigma)$$

$$I = H - \log p, \quad \eta(\iota;\sigma) = e\iota\int_{-\infty}^{-\iota}\varphi\left(\frac{x}{\rho}\right)e^{x}dx = \varphi(-\frac{\iota}{\sigma}) - \exp\left\{\iota\ +\ \frac{1}{2}\sigma^2\right\}\varphi(-\frac{\iota}{\sigma} - \sigma)$$

where σ is the standard deviation of E. For more detailed explanation, see Gapko and Šmíd (Gapko & Šmíd, 2012) or Šmíd and Dufek (Šmíd & Dufek, 2016).

As for the dynamics, we suppose that the vector $(X,\ B,\ H)$ follows a VAR model, i.e.

$$(X_t,\ B_t,\ H_t) = \Gamma U_t + \mathcal{E}_t$$

where Γ is unknown deterministic matrix parameter, $\mathcal{E}_t$ is a Gaussian white noise and the regressors U_t are allowed to include constants, trends, lagged values and exogenous variables. Consequently,

$$Q_t = \varphi(-Y_t) = \varphi\left(-\left[\Gamma^Q U_i + \epsilon_{1,t}\right]\right),$$

$$G_t = \eta(I_t;\sigma) = \eta\left(\Gamma^G U_i + \varepsilon_{2,t};\sigma\right)$$

for some vector parameters Γ^Q and Γ^G and a Gaussian white noise $(\epsilon_1, \varepsilon_2)$.

Once Q_t and G_t are observable, the factors might be got by transformation

$$Y_t = -\varphi^{-1}(Q_t), \qquad I_t = \eta^{-1}(G_t;\sigma)$$

where the (unknown) value of σ might be guessed e.g. from the volatility of a house price index and the average default rate - see the Appendix for details. Consequently, the parameters $\Gamma^{\bullet}$ and the variance matrix of ϵ may be estimated by standard techniques. For the proof of strict monotonicity of $\eta(\bullet;\sigma)$, see Appendix of Gapko and Šmíd (Gapko & Šmíd, 2012).

If, instead of G_t, the charge-off rate

$$L_t = \frac{total\ credit\ loss\ in\ the\ porfolio}{number\ of\ loans\ in\ the\ portfolio}$$

is observed, G_t may be easily computed using the relation $L_t = Q_t G_t$.

3. Data and Empirical Estimation

3.1 Data description

The dataset used consists of four time series, namely residential and commercial mortgage delinquency rates, which are proportions of loans more than 30 days past due (30+) on the total balance, and residential and commercial mortgage charge off rates, which are proportions of charged off loans (net of recoveries) on the average total balance. The dataset was downloaded from the United States Federal Reserve System and thus includes the US nationwide statistics. The time period covered ranges from 1991 to 2016 in a quarterly granularity.

Table 4.1 and Figure 4.1 summarize descriptive statistics and show the time series of the input data. The 30+ delinquency rates were used as proxy metrics for default rates and the charge-off rates represent real losses from the unpaid balance. From the Figure 4.1 it is obvious that the time series are correlated. Also, the recent economic crisis, which started in the US in late 2007 and impacted the US mortgage and real estate markets excessively is visible, as all time series rocketed up to multiples of their preceding values between 2007 and 2010.

Statistic	30+ delinquency rate residential	Charge-off rate residential	30+ delinquency rate commercial	Charge-off rate commercial
Mean value	0.041188	0.004693	0.038408	0.009065
Median	0.023051	0.001584	0.022500	0.002884
Minimum	0.013358	0.000673	0.008500	0.000100
Maximum	0.110150	0.027057	0.120600	0.036297
Standard deviation	0.030813	0.006504	0.031300	0.010536
Variance	0.74810	1.3858	0.81495	1.1623
Skewness	1.1154	1.8399	1.1312	1.1850
Excess kurtosis	-0.332530	2.1346	0.091479	-0.082452
5% percentile	0.010670	0.001345	0.015785	0.0008
95% percentile	0.11055	0.031084	0.10596	0.02177

Table 3.1: Descriptive statistics of input data

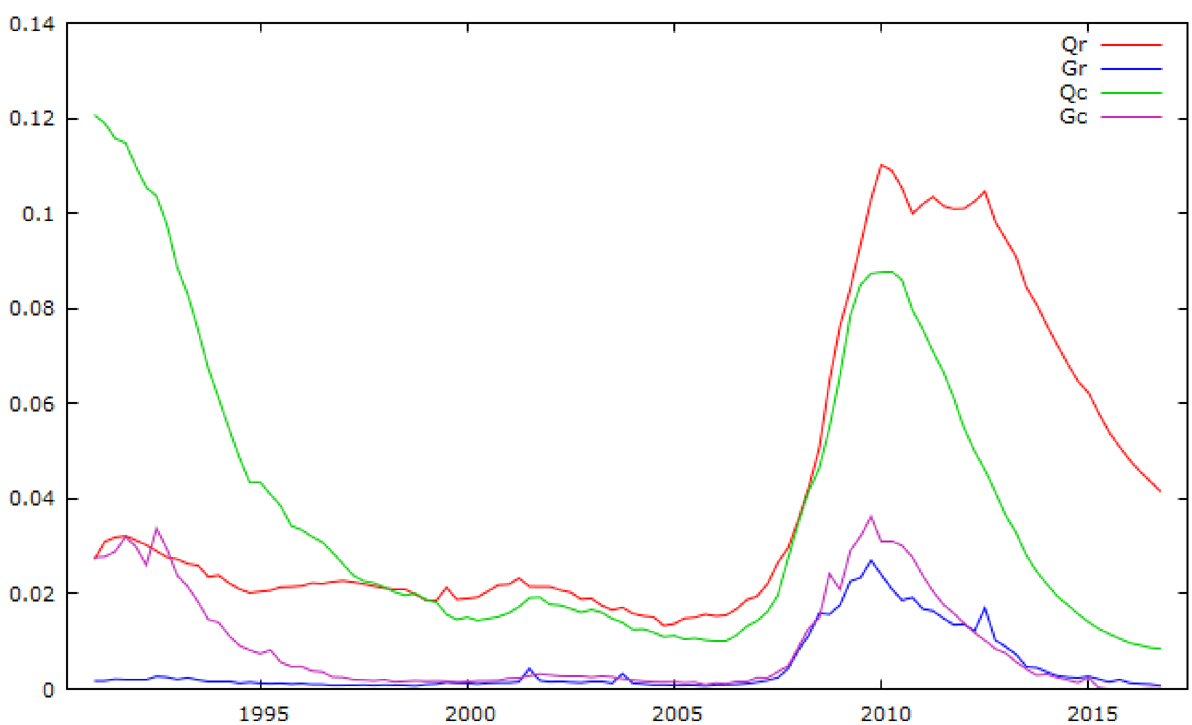

Figure 3.1: Development of the 30+ delinquency rates (Qr = residential, Qc = commercial) and charge-off rates (Gr = residential, Gc = commercial)

3.2 Underlying factors extraction

As the first step in the empirical estimation we extracted the underlying factors by the method described in Section 2. Because the individual factors disappear as a result of application of the central limit theorem, we received four extracted common factors, one or each of the input datasets. The resulting time series of the extracted common factors Y (default rate) and I (loss given default) for both commercial (Yc, Ic) and residential (Yr, Ir) mortgage portfolios are illustrated in Figure 3.2. Similarly to the original input dataset, there is a strong visual correlation, especially between Yr and Yc, and Ir and Ic. This provides us with the basis for the exact estimation of the interconnectedness of the four factors.

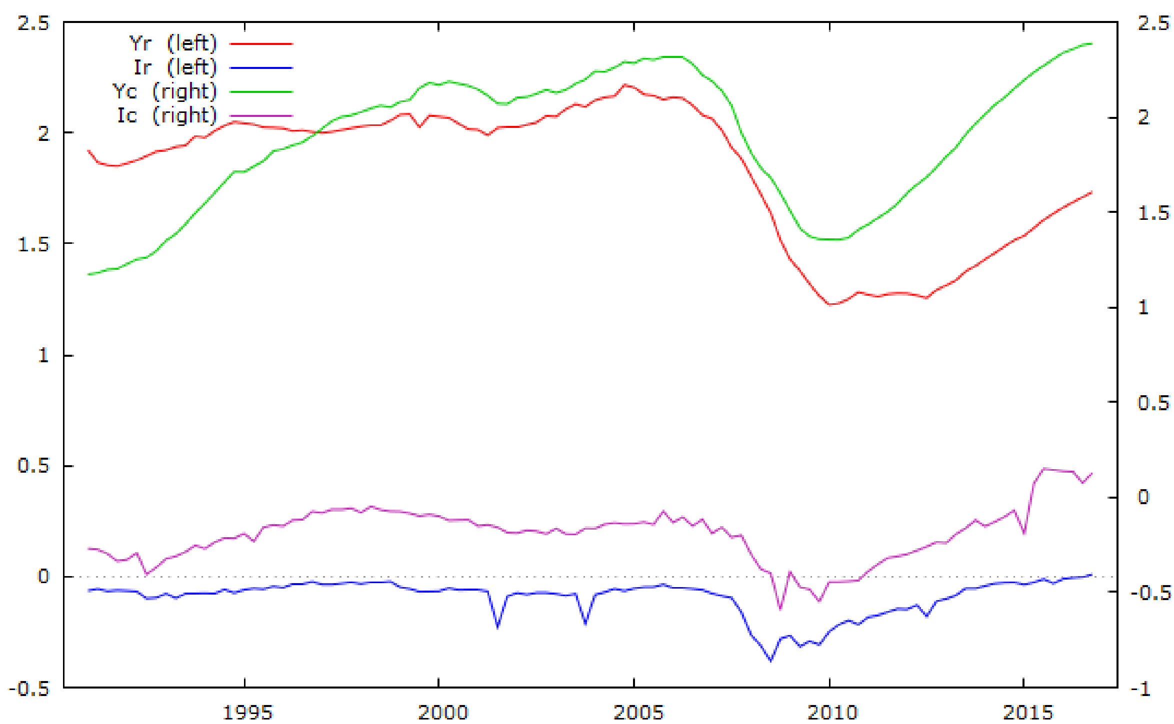

Figure 3.2: The development of the extracted common factors Yr and Ir (left axis), and Yc and Ic (right axis)

3.3 Estimation of interconnectedness of factors and macroeconomic environment

After the extraction of the factors we studied the interconnectedness of the factors and the macroeconomic environment. In our setting, the macroeconomic environment was represented by a set of variables, namely the FED base interest rate, unemployment, GDP, personal income, industrial production and the Case-Shiller HPI index. The macroeconomic variables were chosen to reflect our expectations, i.e. that the default rates are suspected to be driven by the wealth, which was represented by the GDP, unemployment, personal income and industrial production. The HPI index and the interest rate on the other hand are expected to be a driver of property prices, i.e. determine the charge-off rate. In the choice of the macroeconomic factors we were inspired by Pesaran (Pesaran, Schuermann, Treutler, & Weiner, 2003). All time series of macroeconomic variables were obtained from the FED. The analysis was performed on logarithms of all variables and was executed in several steps:

- First, the cointegration was tested for individual factors and the set of macroeconomic variables by the Engle-Granger test
- Second, the cointegration rank were tested for all factors and macroeconomic variables jointly
- Finally, the VECM model for all factors and significant macroeconomic variables was estimated

During the cointegration analysis of individual factors and macroeconomic drivers we focused on the significance of individual variables and also on the simplicity of the cointegrating relationship as the analysis was used in the second step to determine the cointegration rank of all factors and macroeconomic variables jointly.

The cointegration requires the cointegrated variables to be non-stationary. The ADF test of stationarity confirmed on the 90% probability level that all variables were non-stationary. The Engle-Granger cointegration test of individual factors and macroeconomic variables revealed that there exists a strong cointegration among macroeconomic variables and factors Yc, Ic and Ir on a 90% probability level. Only the cointegration among macroeconomic variables and the factor Yr was not confirmed as the stationarity of residuals from the cointegrating regression was rejected. The economic interpretation of the non-significant cointegration here might be that the default rate of individuals was one of the key triggers of particularly the 2007-2009 economic crisis in the United States. The key results of the ADF stationarity tests as well as the results of the Engle-Granger cointegration tests for individual factors is summarized in Table 3.2 and Table 3.3, resp. Details can be found in the Appendix.

Variable	**P-value of the ADF unit root test**	**Stationarity accepted/rejected**
Ic	0.8729	Rejected
Ir	0.5209	Rejected
Yc	0.1895	Rejected
Yr	0.1133	Rejected
U (unemployment)	0.9708	Rejected
PI (personal income)	0.7884	Rejected
IP (industrial production)	0.8330	Rejected
GDP	0.9309	Rejected
HPI	0.9738	Rejected
FEDR (FED interest rate)	0.0048	Accepted

Table 3.2: Results of the ADF stationarity tests

Variable	Ic	Ir	Yc	Yr
Constant	30.951 ***	24.052 ***	22.833 ***	-0.191
U	-0.369 ***	0.022	-0.562 ***	-0.883 ***
PI	-1.815 ***	-1.095 ***	-4.687 ***	-
IP	2.118 ***	2.173 ***	4.047 ***	-
GDP	-2.797 ***	-2.899 ***	-	-

HPI	0.299 ***	0.400 ***	0.660 ***	0.858 ***
FEDR	-	-	-	-0.055 ***
Time	0.030 ***	0.024 ***	0.033 ***	-0.012 ***
R-square	83.7%	76.5%	90.9%	95.4%
Residuals ADF unit root test p-value	0.013	0.012	0.079	0.267

Table 3.3: Results of the Engle-Granger cointegration regressions (significance: * - 90%, ** - 95%, *** - 99%)

In the second step we performed the Johansen cointegration test for all factors and a set of macroeconomic variables resulting from the Engle-Granger tests for individual factors, namely GDP, IP and HPI. The trace and the Lmax tests show that there are three cointegrating relationships in the system as depicted in Table 3.4 (see Appendix for detailed results).

Rank	Eigenvalue	Trace test	p-value	Lmax test	p-value
0	0.389	183.060	0.000	50.237	0.038
1	0.368	132.830	0.000	46.770	0.017
2	0.302	86.057	0.013	36.622	0.054
3	0.222	49.436	0.147	25.567	0.196
4	0.162	23.869	0.460	18.020	0.281
5	0.048	5.849	0.876	5.001	0.895
6	0.008	0.848	0.357	0.848	0.357

Table 3.4: Resuts of the Johansesn cointegration test for all Ic, Ir, Yc, Yr, GDP, IP and HPI.

In the last step, we estimated the VECM model, which allows to include both endogenous and exogenous variables. In our case, we included all variables (i.e. the four factors, GDP, IP and HPI) as endogenous. Furthermore, we also tested for significance a set of exogenous macroeconomic variables. The resulting setting of the estimated VECM is summarized in the Table 3.5. EC_n represents the error correction term of the n-th VECM equation.

Variable	Ic	Ir	Yc	Yr
Constant	-	-	-0.932 *	-4.005 ***
Ic (lag1)	-0.351 ***	-	-	-
Ir (lag1)	-	-0.225 ***	-	-0.200 ***
Yc (lag1)	0.414 **	0.619 ***	0.581 ***	-
HPI (lag1)	-1.149 ***	0.881 ***	0.677 ***	-
GDP (lag1)	-	-1.386 **	-	-
FEDR (lag1)	-0.010 ***	0.008 ***	-0.004 ***	-0.015 ***
EC1	-0.187 ***	0.143 ***	0.108 ***	0.029
EC2	0.320 ***	-0.396 ***	-	0.392 ***
EC3	0.163 **	-	-0.122 ***	0.131 ***

Table 3.5: Results of the VECM estimation (significance: * - 90%, ** - 95%, *** - 99%)

The resulting VECM model confirmed strong interconnectedness of the factors and the macroeconomic environment. The final VECM model can be used to estimate the future value of the factors given the macroeconomic factors and thus predict the credit risk under various macroeconomic scenarios.

4. Conclusion

We constructed a multi-period multi-portfolio dynamic macroeconomic model of credit losses. We estimated the presented model on a large U.S. national portfolio of residential and commercial mortgage loans. The empirical analysis showed that there exists a clear and estimable relationship between the credit risk and the macroeconomic environment. Additionally, we proved that the default rate on the portfolio and the loss given default are not independent, as well as there exists interconnectedness between portfolios. Thus a reasonable model of credit risk has to incorporate the interconnectedness between defaults (represented e.g. by a probability of default) and losses (or, in other words, loss given default) and among risk factors of different portfolios. Finally we demonstrated the possibility of prediction of the credit risk.

The proposed model describes the interconnectedness of the credit risk in the loan book consisting of multiple portfolios with the macroeconomic environment and predicts the credit risk. Additionally, the model, thanks to an inclusion of macroeconomic variables, is capable of estimating potential future development of the portfolio based on different macroeconomic assumptions and therefore can be used as a tool for macroeconomic stress testing. Given all the possible applications, the model can be used as a model of economic capital within a financial institution.

References

Frontczak, R., & Rostek, S. (2015). Modeling loss given default with stochastic collateral. *Economic Modelling*, 44 (2015), pp.162-170.

Frye, J. (2000). Collateral Damage. *Risk*.

Gapko, P., & Šmíd, M. (2012). Dynamic Multi-Factor Credit Risk Model with Fat-Tailed Factors. *Czech Journal of Economics and Finance*, 62(2): 125-140.

Gapko, P., & Šmíd, M. (2012a). Modeling a Distribution of Mortgage Credit Losses. *Ekonomicky casopis*, 1005-1023.

Hamerle, A., Dartsch, A., Jobst, R., & Plank, K. (2011). Integrating macroeconomic risk factors into. *The Journal of Risk Model Validation, vol.5/2*, pp. 3-24.

Hochguertel, S., & Ohlsson, H. (2011). Wealth Mobility and Dynamics Over Entire Individual Working Life Cycles. *ECB Working Paper Series no. 1301*.

Jimenez, G., & Mencia, J. (2009). Modelling the distribution of credit losses with observable and latent factors. *Journal of Empirical Finance*, 235-253.

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journla of Finance 29*, Chapter 12.

Pesaran, M. H., Schuermann, T., Treutler, B.-J., & Weiner, S. M. (2003). Macroeconomic Dynamics and Credit Risk: A Global Perspective. *CESifo Working Paper No. 995*.

Pykhtin, M. V. (2003). Unexpected Recovery Risk. *Risk*, 74-78.

Sommar, P. A., & Shahnazarian, H. (2009, September). Interdependencies between Expected Default Frequency and the Macro Economy. *International Journal of Central Banking*, pp. 83-110.

Šmíd, M. (2015). Model of risk and losses of a multigeneration mortgage portfolio. *10th International Scientific Conference Financial management of firms and financial institutions.* Available at http://ssrn.com.

Šmíd, M., & Dufek, J. (2016). Multi-Period Factor Model of a Loan Portfolio. *Available at SSRN: http://dx.doi.org/10.2139/ssrn.2703884*.

Vasicek, O. A. (1987). *Probability of Loss on Loan Portfolio.* KMV.

Virolainen, K. (2004). Macro Stress Testing with a Macroeconomic Credit Risk Model for Finland. *Bank of Finland Discussion Paper No. 18/2004*.

Witzany, J. (2011). A Two-Factor Model for PD and LGD Correlation. *Bulletin of the Czech Econometric Society*.

Appendix

Determination of σ

Assume that, at time t, the portfolio contains multiple "generations" of loans namely the loans originated at $t-1, t-2, \dots, t-k$ (the loans older than k are no longer present in the portfolio). Assume further that the inflow of fresh loans into the portfolio is constant in time. Finally, assume that all the collaterals securing loans from the generation which started at s have been bought for the same price $\exp\{H_s\}$ and that the price of each of them at t is $\exp\{H_t + (S_t - S_s)\}$ where S is a normal random walk, specific to the loan, with variance θ^2,

Denote G_t the age of a loan randomly chosen at t. Clearly, after k periods, the ratio of the generations within the portfolio is: $(1-q): \dots : (1-q)^{k-1}$, which uniquely determines $\pi_i = \mathbb{P}[G_t = i]$.

Let P_t be the price of a randomly chosen collateral. By the Law of Iterated Variance, we then get

$$\tilde{\sigma}^2 = var(\log P_t|H) = var(\mathbb{E}(\log P_t|G_t, H)|H) + \mathbb{E}(var(\log P_t|G_t, H)|H)$$

$$= var(\mathbb{E}(S_t - S_G|G, H)|H) + \mathbb{E}(var(S_t - S_G|G, H)|H) = \theta^2\, \mathbb{E}G_t = \theta^2 \sum_{i=1}^{k} i\pi_i,$$

Even though the $\mathcal{L}(\log P_t|H)$ is a mixture of normal distributions rather than a normal distribution, it is thin tailed so it will not make a big harm to approximate it by $N(H_t, \tilde{\sigma}^2)$.

Results of the ADF tests and the Engle-Granger cointegration tests for individual factors

Testing for a unit root in Ic

```
Dickey-Fuller test for Ic
sample size 103
unit-root null hypothesis: a = 1

  with constant and trend
  model: (1-L)y = b0 + b1*t + (a-1)*y(-1) + e
  estimated value of (a - 1): -0.0497293
  test statistic: tau_ct(1) = -1.33714
  p-value 0.8729
  1st-order autocorrelation coeff. for e: -0.275
```

Testing for a unit root in Ir

```
Dickey-Fuller test for Ir
sample size 103
unit-root null hypothesis: a = 1

  with constant and trend
  model: (1-L)y = b0 + b1*t + (a-1)*y(-1) + e
  estimated value of (a - 1): -0.100647
  test statistic: tau_ct(1) = -2.13345
  p-value 0.5209
  1st-order autocorrelation coeff. for e: -0.170
```

Testing for a unit root in Yc

```
Augmented Dickey-Fuller test for Yc
including 2 lags of (1-L)Yc
sample size 101
unit-root null hypothesis: a = 1
```

```
  with constant and trend
  model: (1-L)y = b0 + b1*t + (a-1)*y(-1) + ... + e
  estimated value of (a - 1): -0.0184176
  test statistic: tau_ct(1) = -2.82117
  asymptotic p-value 0.1895
  1st-order autocorrelation coeff. for e: -0.032
  lagged differences: F(2, 96) = 134.873 [0.0000]
```

Testing for a unit root in Yr

```
Augmented Dickey-Fuller test for Yr
including 3 lags of (1-L)Yr
sample size 100
unit-root null hypothesis: a = 1

  with constant and trend
  model: (1-L)y = b0 + b1*t + (a-1)*y(-1) + ... + e
  estimated value of (a - 1): -0.0293095
  test statistic: tau_ct(1) = -3.07102
  asymptotic p-value 0.1133
  1st-order autocorrelation coeff. for e: -0.021
  lagged differences: F(3, 94) = 46.857 [0.0000]
```

Testing for a unit root in U

```
Dickey-Fuller test for U
sample size 103
unit-root null hypothesis: a = 1

  with constant and trend
  model: (1-L)y = b0 + b1*t + (a-1)*y(-1) + e
  estimated value of (a - 1): -0.0128755
  test statistic: tau_ct(1) = -0.690306
  p-value 0.9708
  1st-order autocorrelation coeff. for e: 0.675
```

Testing for a unit root in PI

```
Dickey-Fuller test for PI
sample size 103
unit-root null hypothesis: a = 1

  with constant and trend
  model: (1-L)y = b0 + b1*t + (a-1)*y(-1) + e
  estimated value of (a - 1): -0.0428271
  test statistic: tau_ct(1) = -1.5951
  p-value 0.7884
  1st-order autocorrelation coeff. for e: -0.148
```

Testing for a unit root in IP

```
Dickey-Fuller test for IP
sample size 103
unit-root null hypothesis: a = 1

  with constant and trend
  model: (1-L)y = b0 + b1*t + (a-1)*y(-1) + e
  estimated value of (a - 1): -0.0262377
  test statistic: tau_ct(1) = -1.47207
  p-value 0.833
  1st-order autocorrelation coeff. for e: 0.503
```

Testing for a unit root in GDP

```
Dickey-Fuller test for GDP
sample size 103
unit-root null hypothesis: a = 1

  with constant and trend
  model: (1-L)y = b0 + b1*t + (a-1)*y(-1) + e
```

```
  estimated value of (a - 1): -0.0141484
  test statistic: tau_ct(1) = -1.05409
  p-value 0.9309
  1st-order autocorrelation coeff. for e: 0.322
```

Testing for a unit root in HPIr

```
Dickey-Fuller test for HPIr
sample size 103
unit-root null hypothesis: a = 1

  with constant and trend
  model: (1-L)y = b0 + b1*t + (a-1)*y(-1) + e
  estimated value of (a - 1): -0.00810067
  test statistic: tau_ct(1) = -0.646854
  p-value 0.9738
  1st-order autocorrelation coeff. for e: 0.750
```

Testing for a unit root in FEDR

```
Augmented Dickey-Fuller test for FEDR
including 3 lags of (1-L)FEDR
sample size 100
unit-root null hypothesis: a = 1

  with constant and trend
  model: (1-L)y = b0 + b1*t + (a-1)*y(-1) + ... + e
  estimated value of (a - 1): -0.0898696
  test statistic: tau_ct(1) = -4.17634
  asymptotic p-value 0.004789
  1st-order autocorrelation coeff. for e: -0.005
```

Cointegrating regression -
OLS, using observations 1991:1-2016:4 (T = 104)
Dependent variable: Ic

```
             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const      30.9510       3.15466       9.811     3.43e-016 ***
  U          −0.369376     0.0651256    −5.672     1.46e-07  ***
  PI         −1.81483      0.573778     −3.163     0.0021    ***
  IP          2.11787      0.324598      6.525     3.10e-09  ***
  GDP        −2.79685      0.539503     −5.184     1.18e-06  ***
  HPIr        0.299283     0.0806026     3.713     0.0003    ***
  time        0.0298391    0.00283753   10.52      1.04e-017 ***

Mean dependent var  −0.190888   S.D. dependent var   0.144941
Sum squared resid    0.332984   S.E. of regression   0.058590
R-squared            0.846113   Adjusted R-squared   0.836594
Log-likelihood       151.1211   Akaike criterion    −288.2422
Schwarz criterion   −269.7314   Hannan-Quinn        −280.7429
rho                  0.510871   Durbin-Watson        0.977351

Testing for a unit root in uhat

Dickey-Fuller test for uhat
sample size 103
unit-root null hypothesis: a = 1

  model: (1-L)y = (a-1)*y(-1) + e
  estimated value of (a - 1): -0.489129
  test statistic: tau_ct(6) = -5.70672
  p-value 0.01346
  1st-order autocorrelation coeff. for e: -0.047
```

Cointegrating regression -
OLS, using observations 1991:1-2016:4 (T = 104)
Dependent variable: Ir

```
             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
```

```
  const      24.0517       2.08559      11.53     6.89e-020 ***
  U           0.0223668    0.0430554     0.5195   0.6046
  PI         −1.09478      0.379333     −2.886    0.0048    ***
  IP          2.17336      0.214596     10.13     7.11e-017 ***
  GDP        −2.89900      0.356673     −8.128    1.43e-012 ***
  HPIr        0.399876     0.0532875     7.504    2.99e-011 ***
  time        0.0243888    0.00187593   13.00     5.67e-023 ***

Mean dependent var  −0.092010   S.D. dependent var   0.079836
Sum squared resid    0.145538   S.E. of regression   0.038735
R-squared            0.778312   Adjusted R-squared   0.764600
Log-likelihood       194.1593   Akaike criterion    −374.3186
Schwarz criterion   −355.8079   Hannan-Quinn        −366.8193
rho                  0.511374   Durbin-Watson        0.971527

Testing for a unit root in uhat

Dickey-Fuller test for uhat
sample size 103
unit-root null hypothesis: a = 1

  model: (1-L)y = (a-1)*y(-1) + e
  estimated value of (a - 1): -0.488626
  test statistic: tau_ct(6) = -5.73702
  p-value 0.0124
  1st-order autocorrelation coeff. for e: -0.056
```

Cointegrating regression -
OLS, using observations 1991:1-2016:4 (T = 104)
Dependent variable: Yc

```
             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const      22.8334       5.70629       4.001    0.0001    ***
  U          −0.562029     0.0933475    −6.021    3.02e-08  ***
  PI         −4.68746      0.800614     −5.855    6.37e-08  ***
  IP          4.04688      0.452070      8.952    2.27e-014 ***
  HPIr        0.659988     0.112513      5.866    6.06e-08  ***
  time        0.0334555    0.00460100    7.271    8.78e-011 ***

Mean dependent var   1.883028   S.D. dependent var   0.353531
Sum squared resid    1.108573   S.E. of regression   0.106358
R-squared            0.913886   Adjusted R-squared   0.909493
Log-likelihood       88.57890   Akaike criterion    −165.1578
Schwarz criterion   −149.2915   Hannan-Quinn        −158.7299
rho                  0.859357   Durbin-Watson        0.256658

Step 7: testing for a unit root in uhat

Augmented Dickey-Fuller test for uhat
including 2 lags of (1-L)uhat
sample size 101
unit-root null hypothesis: a = 1

  model: (1-L)y = (a-1)*y(-1) + ... + e
  estimated value of (a - 1): -0.233003
  test statistic: tau_ct(5) = -4.53415
  asymptotic p-value 0.07904
  1st-order autocorrelation coeff. for e: -0.018
  lagged differences: F(2, 98) = 8.805 [0.0003]
```

Cointegrating regression -
OLS, using observations 1991:1-2016:4 (T = 104)
Dependent variable: Yr

```
             coefficient   std. error    t-ratio     p-value
  -----------------------------------------------------------
  const      −0.190922     0.307739      −0.6204   0.5364
  U          −0.882586     0.0421590    −20.93     3.85e-038 ***
  HPIr        0.858367     0.0553383     15.51     3.00e-028 ***
  FEDR       −0.0552038    0.00661670    −8.343    4.39e-013 ***
```

```
  time       -0.0116112    0.000453633   -25.60     1.92e-045 ***

Mean dependent var   1.825992   S.D. dependent var   0.310982
Sum squared resid    0.436184   S.E. of regression   0.066377
R-squared            0.956211   Adjusted R-squared   0.954442
Log-likelihood       137.0827   Akaike criterion    -264.1654
Schwarz criterion   -250.9435   Hannan-Quinn        -258.8088
rho                  0.868036   Durbin-Watson        0.254628

Testing for a unit root in uhat

Augmented Dickey-Fuller test for uhat
including 3 lags of (1-L)uhat
sample size 100
unit-root null hypothesis: a = 1

  model: (1-L)y = (a-1)*y(-1) + ... + e
  estimated value of (a - 1): -0.195126
  test statistic: tau_ct(4) = -3.64373
  asymptotic p-value 0.2673
  1st-order autocorrelation coeff. for e: 0.050
  lagged differences: F(3, 96) = 2.247 [0.0877]
```

Results of the Johansen cointegration test

Johansen test: Variables Yr, Yc, Ic, Ir, GDP, IP, HPI

```
Number of equations = 7
Lag order = 2
Estimation period: 1991:3 - 2016:4 (T = 102)
Case 5: Unrestricted trend and constant

Log-likelihood = 2376.42 (including constant term: 2086.95)

Rank Eigenvalue Trace test  p-value  Lmax test  p-value
   0    0.38892     183.06 [0.0000]     50.237 [0.0381]
   1    0.36779     132.83 [0.0003]     46.770 [0.0166]
   2    0.30165     86.057 [0.0127]     36.622 [0.0537]
   3    0.22171     49.436 [0.1473]     25.567 [0.1958]
   4    0.16194     23.869 [0.4597]     18.020 [0.2807]
   5   0.047844     5.8487 [0.8762]     5.0007 [0.8952]
   6  0.0082792    0.84799 [0.3571]    0.84799 [0.3571]

Corrected for sample size (df = 86)
Rank Trace test p-value
   0     183.06 [0.0000]
   1     132.83 [0.0006]
   2     86.057 [0.0175]
   3     49.436 [0.1624]
   4     23.869 [0.4691]
   5     5.8487 [0.8767]
   6    0.84799 [0.3636]
```