

Modeling Rating Transition Matrices for Wholesale Loan Portfolios with Stata

Christopher F Baum Alper Corlu Soner Tunay

Boston College / DIW Berlin Risk Analytics, Citizens Financial Group

Stata Conference 2016, Chicago

Modeling asset quality rating transition matrices

- Evaluation of the response of commercial banks' loan portfolios to stressful economic conditions is mandated in the US by Dodd–Frank and the Fed's CCAR stress testing scenarios imposed on SIFIs.
- In this paper, we focus on how the asset quality ratings (AQRs) of wholesale loans may vary in response to changes in the macroeconomic environment.
- A number of approaches have been developed to model the AQR transition matrices of wholesale loans' asset quality rating measures.

Modeling asset quality rating transition matrices

- Evaluation of the response of commercial banks' loan portfolios to stressful economic conditions is mandated in the US by Dodd–Frank and the Fed's CCAR stress testing scenarios imposed on SIFIs.
- In this paper, we focus on how the asset quality ratings (AQRs) of wholesale loans may vary in response to changes in the macroeconomic environment.
- A number of approaches have been developed to model the AQR transition matrices of wholesale loans' asset quality rating measures.

Modeling asset quality rating transition matrices

- Evaluation of the response of commercial banks' loan portfolios to stressful economic conditions is mandated in the US by Dodd–Frank and the Fed's CCAR stress testing scenarios imposed on SIFIs.
- In this paper, we focus on how the asset quality ratings (AQRs) of wholesale loans may vary in response to changes in the macroeconomic environment.
- A number of approaches have been developed to model the AQR transition matrices of wholesale loans' asset quality rating measures.

- A commonly used technique in the financial industry is the single index approach, where creditworthiness is modeled as responding to a systematic factor, Z_t .
- Z_t , meant to reflect the credit cycle, is expressed as a standard normal variable, explaining the deviation of the transition matrix from the average transition matrix.
- The inherent symmetry in this approach is a weakness, as in times of stress the distribution of transitions may become skewed, and larger transitions may be more frequent.
- The strongest objection to this approach: the default state is treated like any other column of the matrix, whereas it plays a crucial role in forecasting losses from the portfolio.
- We have recently implemented the single-index approach using Mata's `optimize` functions.

- A commonly used technique in the financial industry is the single index approach, where creditworthiness is modeled as responding to a systematic factor, Z_t .
- Z_t , meant to reflect the credit cycle, is expressed as a standard normal variable, explaining the deviation of the transition matrix from the average transition matrix.
- The inherent symmetry in this approach is a weakness, as in times of stress the distribution of transitions may become skewed, and larger transitions may be more frequent.
- The strongest objection to this approach: the default state is treated like any other column of the matrix, whereas it plays a crucial role in forecasting losses from the portfolio.
- We have recently implemented the single-index approach using Mata's `optimize` functions.

- A commonly used technique in the financial industry is the single index approach, where creditworthiness is modeled as responding to a systematic factor, Z_t .
- Z_t , meant to reflect the credit cycle, is expressed as a standard normal variable, explaining the deviation of the transition matrix from the average transition matrix.
- The inherent symmetry in this approach is a weakness, as in times of stress the distribution of transitions may become skewed, and larger transitions may be more frequent.
- The strongest objection to this approach: the default state is treated like any other column of the matrix, whereas it plays a crucial role in forecasting losses from the portfolio.
- We have recently implemented the single-index approach using Mata's `optimize` functions.

- A commonly used technique in the financial industry is the single index approach, where creditworthiness is modeled as responding to a systematic factor, Z_t .
- Z_t , meant to reflect the credit cycle, is expressed as a standard normal variable, explaining the deviation of the transition matrix from the average transition matrix.
- The inherent symmetry in this approach is a weakness, as in times of stress the distribution of transitions may become skewed, and larger transitions may be more frequent.
- The strongest objection to this approach: the default state is treated like any other column of the matrix, whereas it plays a crucial role in forecasting losses from the portfolio.
- We have recently implemented the single-index approach using Mata's `optimize` functions.

- A commonly used technique in the financial industry is the single index approach, where creditworthiness is modeled as responding to a systematic factor, Z_t .
- Z_t , meant to reflect the credit cycle, is expressed as a standard normal variable, explaining the deviation of the transition matrix from the average transition matrix.
- The inherent symmetry in this approach is a weakness, as in times of stress the distribution of transitions may become skewed, and larger transitions may be more frequent.
- The strongest objection to this approach: the default state is treated like any other column of the matrix, whereas it plays a crucial role in forecasting losses from the portfolio.
- We have recently implemented the single-index approach using Mata's `optimize` functions.

Advantages of Stata-based risk modeling

- Stata's facilities for data management and econometric modeling have made it the tool of choice for many of the analytical tasks encountered in CFG's retail and wholesale risk analysis.
- Our first approach to risk analysis of wholesale loans is based on the fractional logit model, available for some time as a GLM, but now explicitly supported by `fracreg logit`.
- Our fractional logit model estimates an entire set of models for different rating classes in a single step.
- Our second approach to risk analysis of wholesale loans is based on the SUR model(`sureg`), well supported in Stata for both estimation and postestimation tasks.

Advantages of Stata-based risk modeling

- Stata's facilities for data management and econometric modeling have made it the tool of choice for many of the analytical tasks encountered in CFG's retail and wholesale risk analysis.
- Our first approach to risk analysis of wholesale loans is based on the fractional logit model, available for some time as a GLM, but now explicitly supported by `fracreg logit`.
- Our fractional logit model estimates an entire set of models for different rating classes in a single step.
- Our second approach to risk analysis of wholesale loans is based on the SUR model(`sureg`), well supported in Stata for both estimation and postestimation tasks.

Advantages of Stata-based risk modeling

- Stata's facilities for data management and econometric modeling have made it the tool of choice for many of the analytical tasks encountered in CFG's retail and wholesale risk analysis.
- Our first approach to risk analysis of wholesale loans is based on the fractional logit model, available for some time as a GLM, but now explicitly supported by `fracreg logit`.
- Our fractional logit model estimates an entire set of models for different rating classes in a single step.
- Our second approach to risk analysis of wholesale loans is based on the SUR model(`sureg`), well supported in Stata for both estimation and postestimation tasks.

Advantages of Stata-based risk modeling

- Stata's facilities for data management and econometric modeling have made it the tool of choice for many of the analytical tasks encountered in CFG's retail and wholesale risk analysis.
- Our first approach to risk analysis of wholesale loans is based on the fractional logit model, available for some time as a GLM, but now explicitly supported by `fracreg logit`.
- Our fractional logit model estimates an entire set of models for different rating classes in a single step.
- Our second approach to risk analysis of wholesale loans is based on the SUR model(`sureg`), well supported in Stata for both estimation and postestimation tasks.

Our AQR modeling approach

- We depart from the single-index approach (or its multi-index generalization) to propose a more flexible method of modeling AQR transitions.
- Our approach involves the modeling of the most likely events: (i) no change in asset quality, (ii) one-notch changes up or down on the asset quality scale, (iii) transition to default.
- As historical transition matrices are dominated by the diagonal and adjacent diagonals, this approach should be able to capture a large fraction of experienced ratings changes.

Our AQR modeling approach

- We depart from the single-index approach (or its multi-index generalization) to propose a more flexible method of modeling AQR transitions.
- Our approach involves the modeling of the most likely events: (i) no change in asset quality, (ii) one-notch changes up or down on the asset quality scale, (iii) transition to default.
- As historical transition matrices are dominated by the diagonal and adjacent diagonals, this approach should be able to capture a large fraction of experienced ratings changes.

Our AQR modeling approach

- We depart from the single-index approach (or its multi-index generalization) to propose a more flexible method of modeling AQR transitions.
- Our approach involves the modeling of the most likely events: (i) no change in asset quality, (ii) one-notch changes up or down on the asset quality scale, (iii) transition to default.
- As historical transition matrices are dominated by the diagonal and adjacent diagonals, this approach should be able to capture a large fraction of experienced ratings changes.

Modeling the most likely transitions

This approach is implemented, using time series data on 16 asset quality ratings (AAA, AA+, AA, AA-...), with a set of *fractional logit* models of the probabilities of four events:

- 1 AQ unchanged from period $t - 1$ to t
- 2 AQ increased by one rating from period $t - 1$ to t
- 3 AQ decreased by one rating from period $t - 1$ to t
- 4 AQ transition to default in period t

The remaining transition probabilities from the period $t - 1$ transition matrix are mechanically adjusted to meet the constraint that the probabilities of events (1)–(4) plus remaining probabilities must sum to one for each AQ rating. This implies that the AQ transition matrices evolve dynamically in a relatively unconstrained manner.

Modeling the most likely transitions

This approach is implemented, using time series data on 16 asset quality ratings (AAA, AA+, AA, AA-...), with a set of *fractional logit* models of the probabilities of four events:

- 1 AQ unchanged from period $t - 1$ to t
- 2 AQ increased by one rating from period $t - 1$ to t
- 3 AQ decreased by one rating from period $t - 1$ to t
- 4 AQ transition to default in period t

The remaining transition probabilities from the period $t - 1$ transition matrix are mechanically adjusted to meet the constraint that the probabilities of events (1)–(4) plus remaining probabilities must sum to one for each AQ rating. This implies that the AQ transition matrices evolve dynamically in a relatively unconstrained manner.

Modeling the most likely transitions

This approach is implemented, using time series data on 16 asset quality ratings (AAA, AA+, AA, AA-...), with a set of *fractional logit* models of the probabilities of four events:

- 1 AQ unchanged from period $t - 1$ to t
- 2 AQ increased by one rating from period $t - 1$ to t
- 3 AQ decreased by one rating from period $t - 1$ to t
- 4 AQ transition to default in period t

The remaining transition probabilities from the period $t - 1$ transition matrix are mechanically adjusted to meet the constraint that the probabilities of events (1)–(4) plus remaining probabilities must sum to one for each AQ rating. This implies that the AQ transition matrices evolve dynamically in a relatively unconstrained manner.

Modeling the most likely transitions

This approach is implemented, using time series data on 16 asset quality ratings (AAA, AA+, AA, AA-...), with a set of *fractional logit* models of the probabilities of four events:

- 1 AQ unchanged from period $t - 1$ to t
- 2 AQ increased by one rating from period $t - 1$ to t
- 3 AQ decreased by one rating from period $t - 1$ to t
- 4 AQ transition to default in period t

The remaining transition probabilities from the period $t - 1$ transition matrix are mechanically adjusted to meet the constraint that the probabilities of events (1)–(4) plus remaining probabilities must sum to one for each AQ rating. This implies that the AQ transition matrices evolve dynamically in a relatively unconstrained manner.

Modeling the most likely transitions

This approach is implemented, using time series data on 16 asset quality ratings (AAA, AA+, AA, AA-...), with a set of *fractional logit* models of the probabilities of four events:

- ① AQ unchanged from period $t - 1$ to t
- ② AQ increased by one rating from period $t - 1$ to t
- ③ AQ decreased by one rating from period $t - 1$ to t
- ④ AQ transition to default in period t

The remaining transition probabilities from the period $t - 1$ transition matrix are mechanically adjusted to meet the constraint that the probabilities of events (1)–(4) plus remaining probabilities must sum to one for each AQ rating. This implies that the AQ transition matrices evolve dynamically in a relatively unconstrained manner.

Modeling the most likely transitions

This approach is implemented, using time series data on 16 asset quality ratings (AAA, AA+, AA, AA-...), with a set of *fractional logit* models of the probabilities of four events:

- ① AQ unchanged from period $t - 1$ to t
- ② AQ increased by one rating from period $t - 1$ to t
- ③ AQ decreased by one rating from period $t - 1$ to t
- ④ AQ transition to default in period t

The remaining transition probabilities from the period $t - 1$ transition matrix are mechanically adjusted to meet the constraint that the probabilities of events (1)–(4) plus remaining probabilities must sum to one for each AQ rating. This implies that the AQ transition matrices evolve dynamically in a relatively unconstrained manner.

Modeling the most likely transitions

This approach is implemented, using time series data on 16 asset quality ratings (AAA, AA+, AA, AA-...), with a set of *fractional logit* models of the probabilities of four events:

- ① AQ unchanged from period $t - 1$ to t
- ② AQ increased by one rating from period $t - 1$ to t
- ③ AQ decreased by one rating from period $t - 1$ to t
- ④ AQ transition to default in period t

The remaining transition probabilities from the period $t - 1$ transition matrix are mechanically adjusted to meet the constraint that the probabilities of events (1)–(4) plus remaining probabilities must sum to one for each AQ rating. This implies that the AQ transition matrices evolve dynamically in a relatively unconstrained manner.

The fractional logit model

- The fractional logit model of Papke and Wooldridge (*J. Applied Econometrics*, 1996) is appropriate when the response variable is a proportion which may take on the values 0 and 1.
- Proportions data should not be modeled using linear regression, as it does not respect the bounds of 0 and 1, and thus can produce predictions outside the (0,1) interval.
- In contrast to the Tobit model, which separately models the likelihood that the response is 0 (or some other limiting value), the fractional logit produces a single set of marginal effects which show the effects of each explanatory variable on the proportion.
- The fractional logit model produces predictions on the same scale as the original response variable, avoiding retransformation bias.

The fractional logit model

- The fractional logit model of Papke and Wooldridge (*J. Applied Econometrics*, 1996) is appropriate when the response variable is a proportion which may take on the values 0 and 1.
- Proportions data should not be modeled using linear regression, as it does not respect the bounds of 0 and 1, and thus can produce predictions outside the (0,1) interval.
- In contrast to the Tobit model, which separately models the likelihood that the response is 0 (or some other limiting value), the fractional logit produces a single set of marginal effects which show the effects of each explanatory variable on the proportion.
- The fractional logit model produces predictions on the same scale as the original response variable, avoiding retransformation bias.

The fractional logit model

- The fractional logit model of Papke and Wooldridge (*J. Applied Econometrics*, 1996) is appropriate when the response variable is a proportion which may take on the values 0 and 1.
- Proportions data should not be modeled using linear regression, as it does not respect the bounds of 0 and 1, and thus can produce predictions outside the (0,1) interval.
- In contrast to the Tobit model, which separately models the likelihood that the response is 0 (or some other limiting value), the fractional logit produces a single set of marginal effects which show the effects of each explanatory variable on the proportion.
- The fractional logit model produces predictions on the same scale as the original response variable, avoiding retransformation bias.

The fractional logit model

- The fractional logit model of Papke and Wooldridge (*J. Applied Econometrics*, 1996) is appropriate when the response variable is a proportion which may take on the values 0 and 1.
- Proportions data should not be modeled using linear regression, as it does not respect the bounds of 0 and 1, and thus can produce predictions outside the (0,1) interval.
- In contrast to the Tobit model, which separately models the likelihood that the response is 0 (or some other limiting value), the fractional logit produces a single set of marginal effects which show the effects of each explanatory variable on the proportion.
- The fractional logit model produces predictions on the same scale as the original response variable, avoiding retransformation bias.

- A stylized fact of transition matrix modeling is that the probability of an unchanged AQ rating (item (1) above) is substantial.
- Our fractional logit model of the probability of 'staying put' takes into account the probabilities of transitions from each of the non-default AQ ratings, augmented with relevant macroeconomic factors.
- In the estimated model, we also found that interactions of the macro factors with transition probabilities, as well as with each other, are very relevant.
- The model, fit to quarterly data for a portfolio of C&I loans from 2000–2012 over 15 non-default AQ classes, yielded an $R^2 = 0.604$.

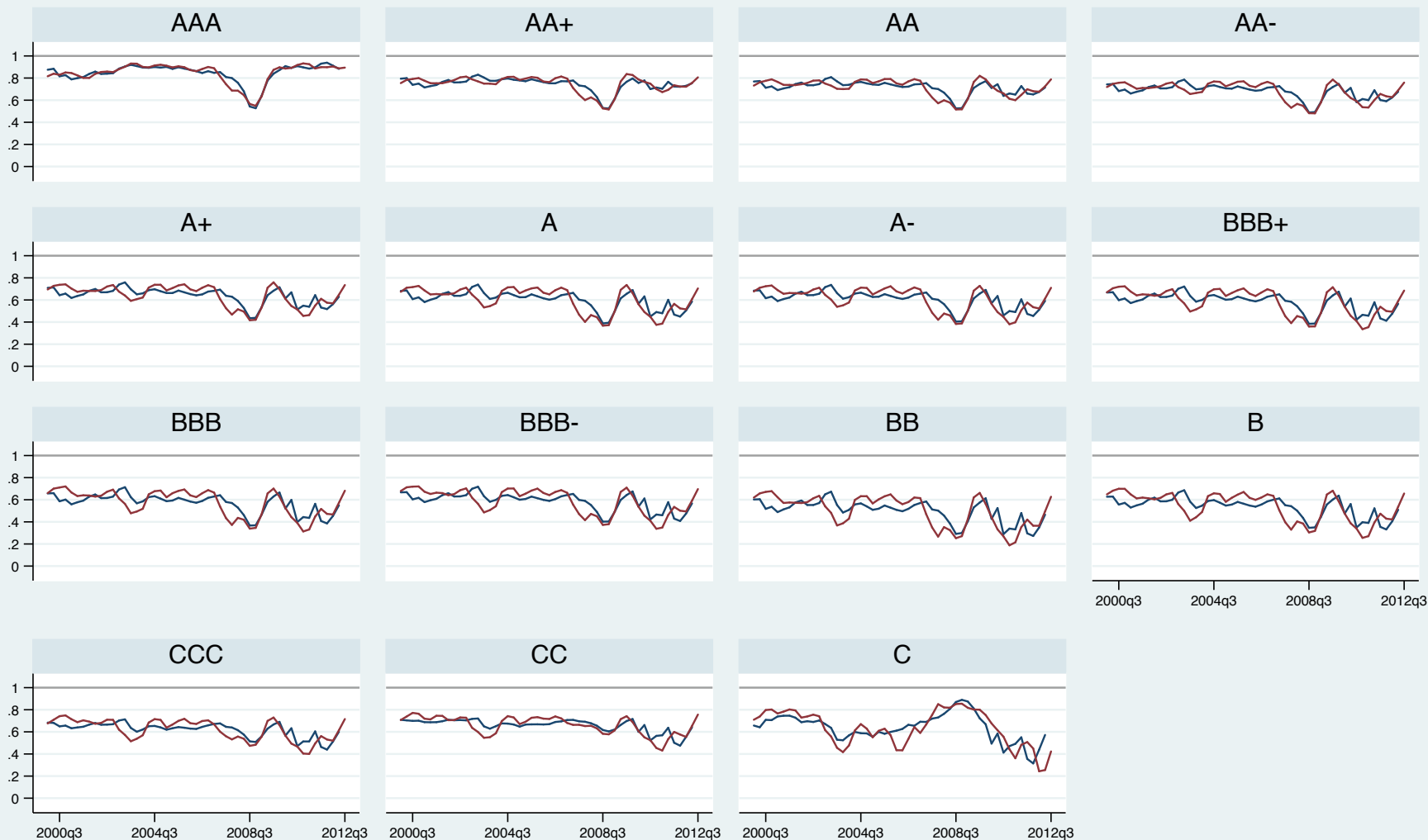
- A stylized fact of transition matrix modeling is that the probability of an unchanged AQ rating (item (1) above) is substantial.
- Our fractional logit model of the probability of 'staying put' takes into account the probabilities of transitions from each of the non-default AQ ratings, augmented with relevant macroeconomic factors.
- In the estimated model, we also found that interactions of the macro factors with transition probabilities, as well as with each other, are very relevant.
- The model, fit to quarterly data for a portfolio of C&I loans from 2000–2012 over 15 non-default AQ classes, yielded an $R^2 = 0.604$.

- A stylized fact of transition matrix modeling is that the probability of an unchanged AQ rating (item (1) above) is substantial.
- Our fractional logit model of the probability of 'staying put' takes into account the probabilities of transitions from each of the non-default AQ ratings, augmented with relevant macroeconomic factors.
- In the estimated model, we also found that interactions of the macro factors with transition probabilities, as well as with each other, are very relevant.
- The model, fit to quarterly data for a portfolio of C&I loans from 2000–2012 over 15 non-default AQ classes, yielded an $R^2 = 0.604$.

- A stylized fact of transition matrix modeling is that the probability of an unchanged AQ rating (item (1) above) is substantial.
- Our fractional logit model of the probability of 'staying put' takes into account the probabilities of transitions from each of the non-default AQ ratings, augmented with relevant macroeconomic factors.
- In the estimated model, we also found that interactions of the macro factors with transition probabilities, as well as with each other, are very relevant.
- The model, fit to quarterly data for a portfolio of C&I loans from 2000–2012 over 15 non-default AQ classes, yielded an $R^2 = 0.604$.

Predicted Staying the Same Probability

blue: predicted series using gdpqgr and neg/pos unempDunemp



Graphs by fr

- After ‘staying put’, the most likely outcomes from one period to the next are one-notch upgrades or downgrades in the AQ rating: items (2) and (3) above.
- These are modeled with separate fractional logits of the same specification, taking transitions from each AQ rating as explanatory factors.
- The same macroeconomic factors are included, interacted with the AQ rating classes.
- The one-notch upgrade model yielded an $R^2 = 0.523$ while the one-notch downgrade model yielded an $R^2 = 0.656$.

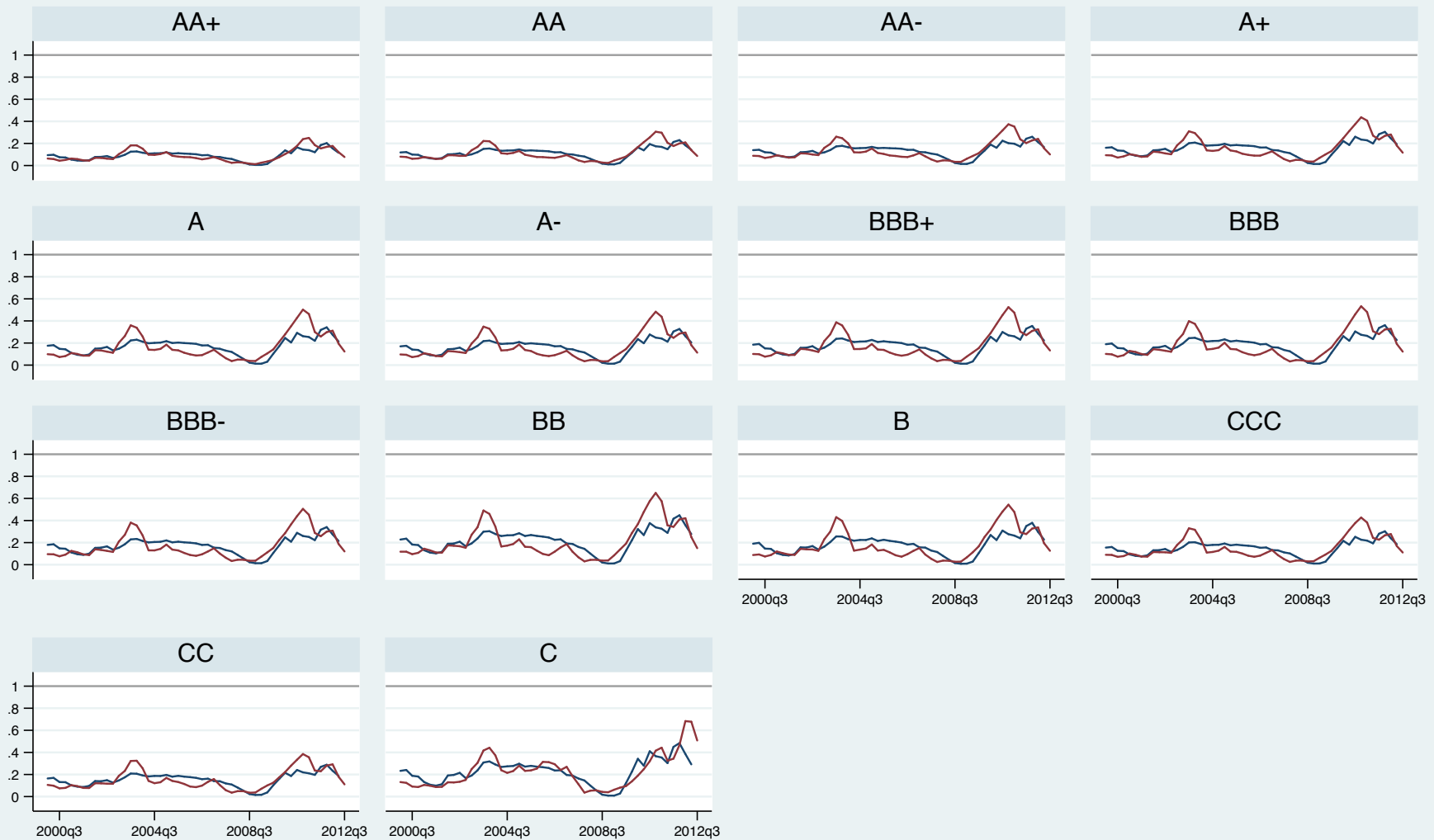
- After ‘staying put’, the most likely outcomes from one period to the next are one-notch upgrades or downgrades in the AQ rating: items (2) and (3) above.
- These are modeled with separate fractional logits of the same specification, taking transitions from each AQ rating as explanatory factors.
- The same macroeconomic factors are included, interacted with the AQ rating classes.
- The one-notch upgrade model yielded an $R^2 = 0.523$ while the one-notch downgrade model yielded an $R^2 = 0.656$.

- After ‘staying put’, the most likely outcomes from one period to the next are one-notch upgrades or downgrades in the AQ rating: items (2) and (3) above.
- These are modeled with separate fractional logits of the same specification, taking transitions from each AQ rating as explanatory factors.
- The same macroeconomic factors are included, interacted with the AQ rating classes.
- The one-notch upgrade model yielded an $R^2 = 0.523$ while the one-notch downgrade model yielded an $R^2 = 0.656$.

- After ‘staying put’, the most likely outcomes from one period to the next are one-notch upgrades or downgrades in the AQ rating: items (2) and (3) above.
- These are modeled with separate fractional logits of the same specification, taking transitions from each AQ rating as explanatory factors.
- The same macroeconomic factors are included, interacted with the AQ rating classes.
- The one-notch upgrade model yielded an $R^2 = 0.523$ while the one-notch downgrade model yielded an $R^2 = 0.656$.

Predicted Upgrades

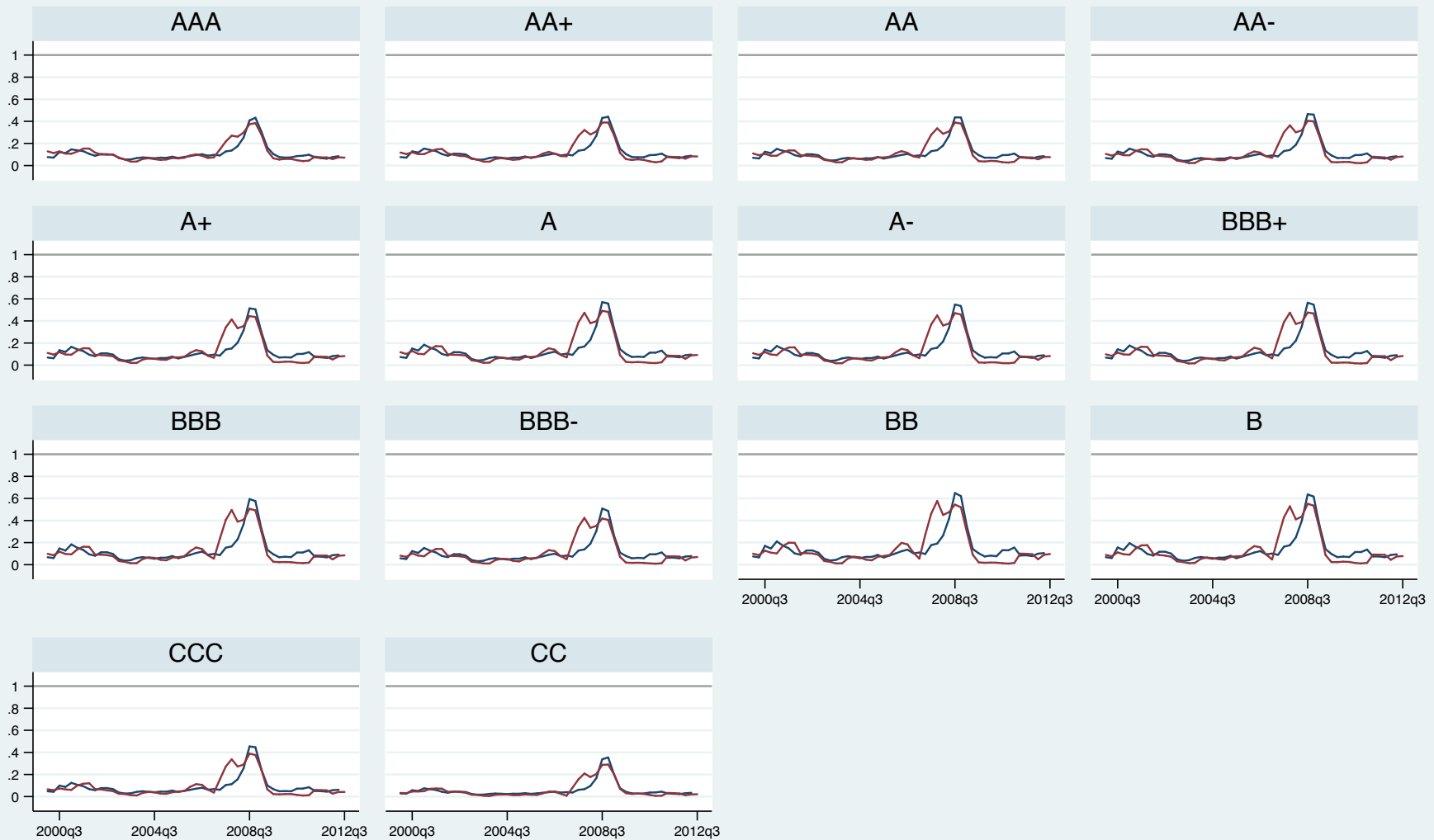
blue: predicted series using unempDunemp and gdpqgr



Graphs by fr

Predicted Downgrades

blue: predicted series using unempDunemp and gdpqgr



Graphs by fr

- The last item to be explicitly modeled in this approach is the transition to default, item (4) above.
- The fractional logit model of the probability of default takes transitions from each AQ ratio as explanatory factors.
- Macroeconomic factors are included, and interacted with the AQ ratio classes.
- The default model yielded an $R^2 = 0.987$.

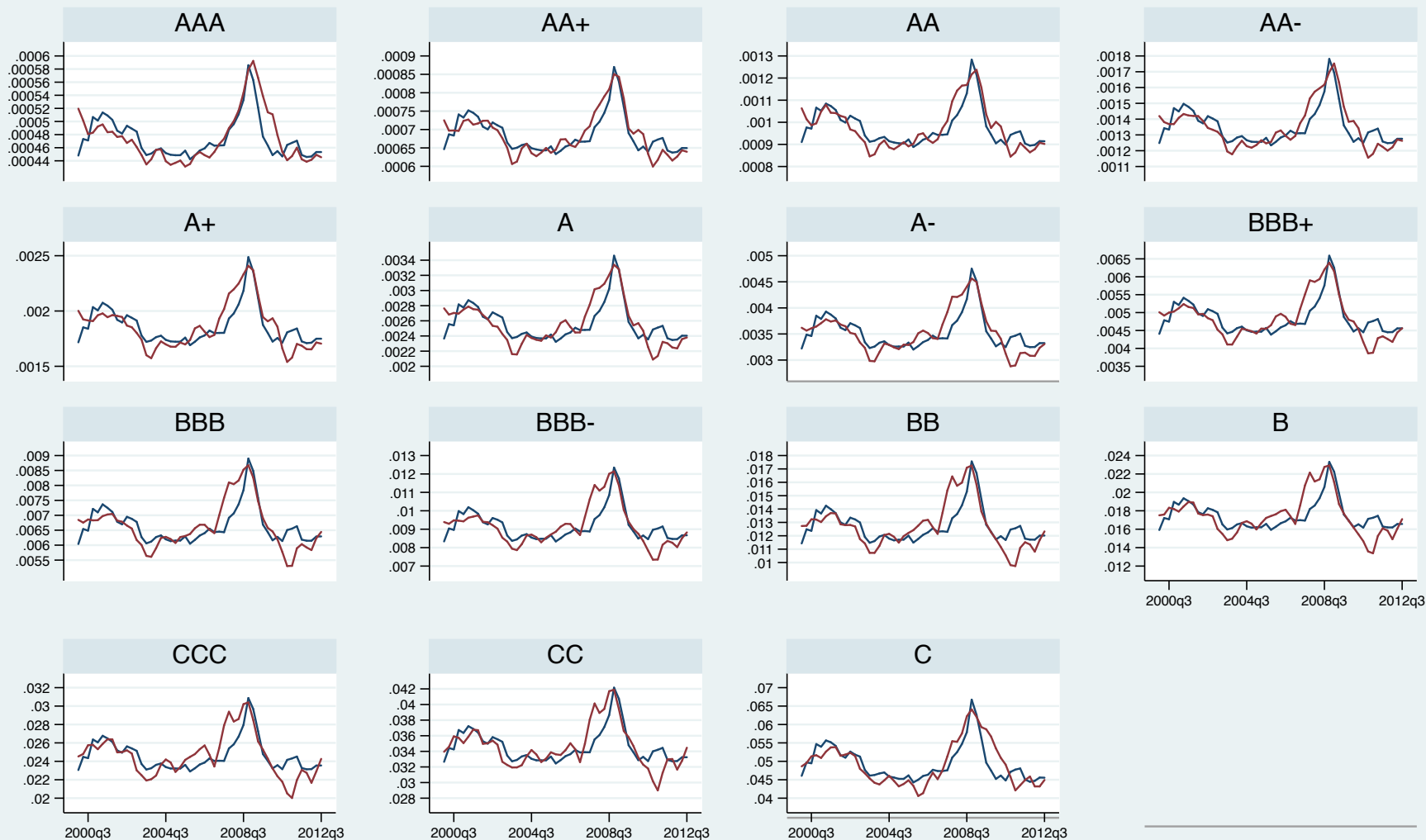
- The last item to be explicitly modeled in this approach is the transition to default, item (4) above.
- The fractional logit model of the probability of default takes transitions from each AQ ratio as explanatory factors.
- Macroeconomic factors are included, and interacted with the AQ ratio classes.
- The default model yielded an $R^2 = 0.987$.

- The last item to be explicitly modeled in this approach is the transition to default, item (4) above.
- The fractional logit model of the probability of default takes transitions from each AQ ratio as explanatory factors.
- Macroeconomic factors are included, and interacted with the AQ ratio classes.
- The default model yielded an $R^2 = 0.987$.

- The last item to be explicitly modeled in this approach is the transition to default, item (4) above.
- The fractional logit model of the probability of default takes transitions from each AQ ratio as explanatory factors.
- Macroeconomic factors are included, and interacted with the AQ ratio classes.
- The default model yielded an $R^2 = 0.987$.

Predicted Defaults

blue: predicted series using gdpqgr and corpprofit

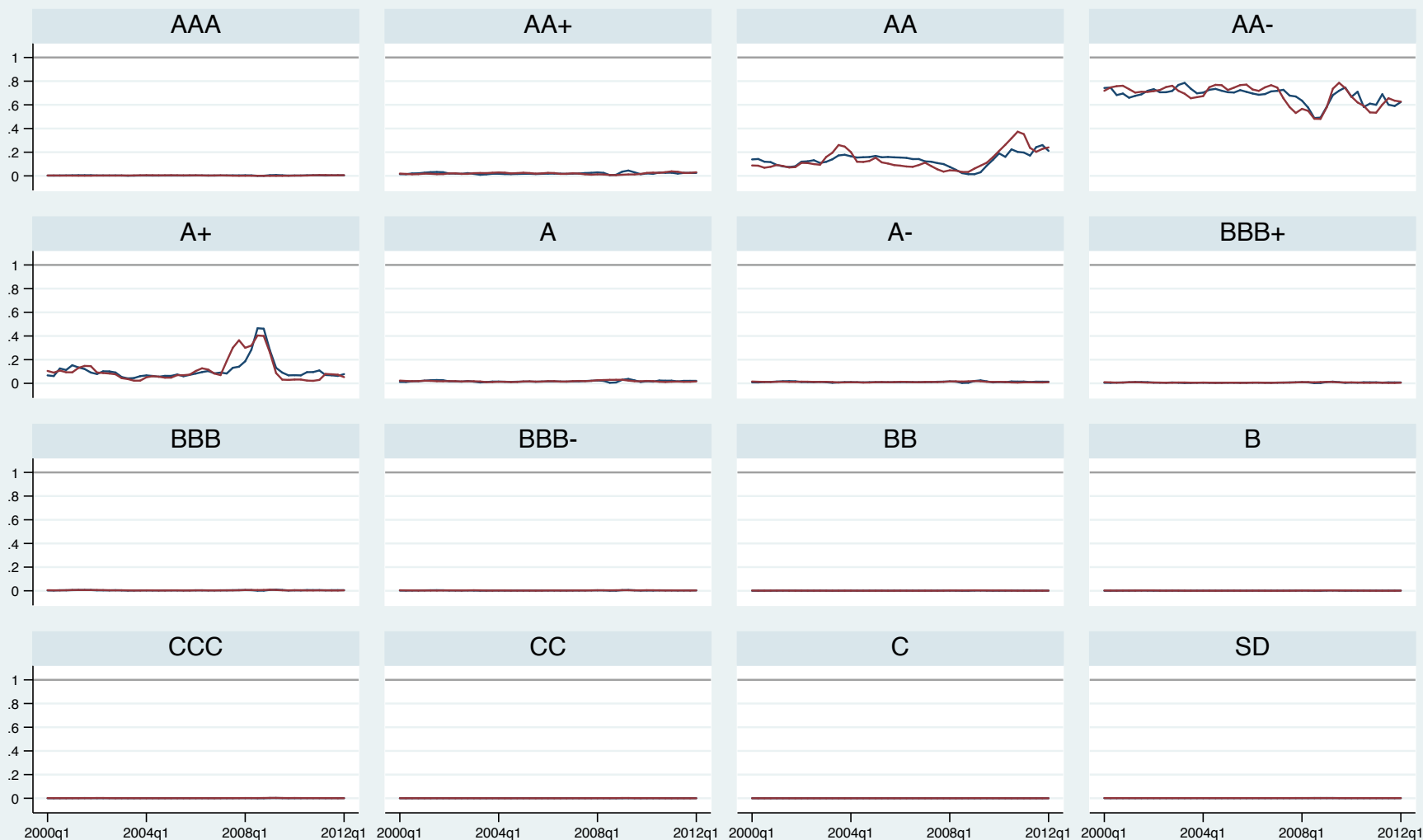


Graphs by fr

We can also visualize the predictions of the model by asset quality rating, in terms of the probability that a loan of a given rating is likely to stay in that rating, migrate to a different rating, or default. For example:

Predicted transitions from AA-

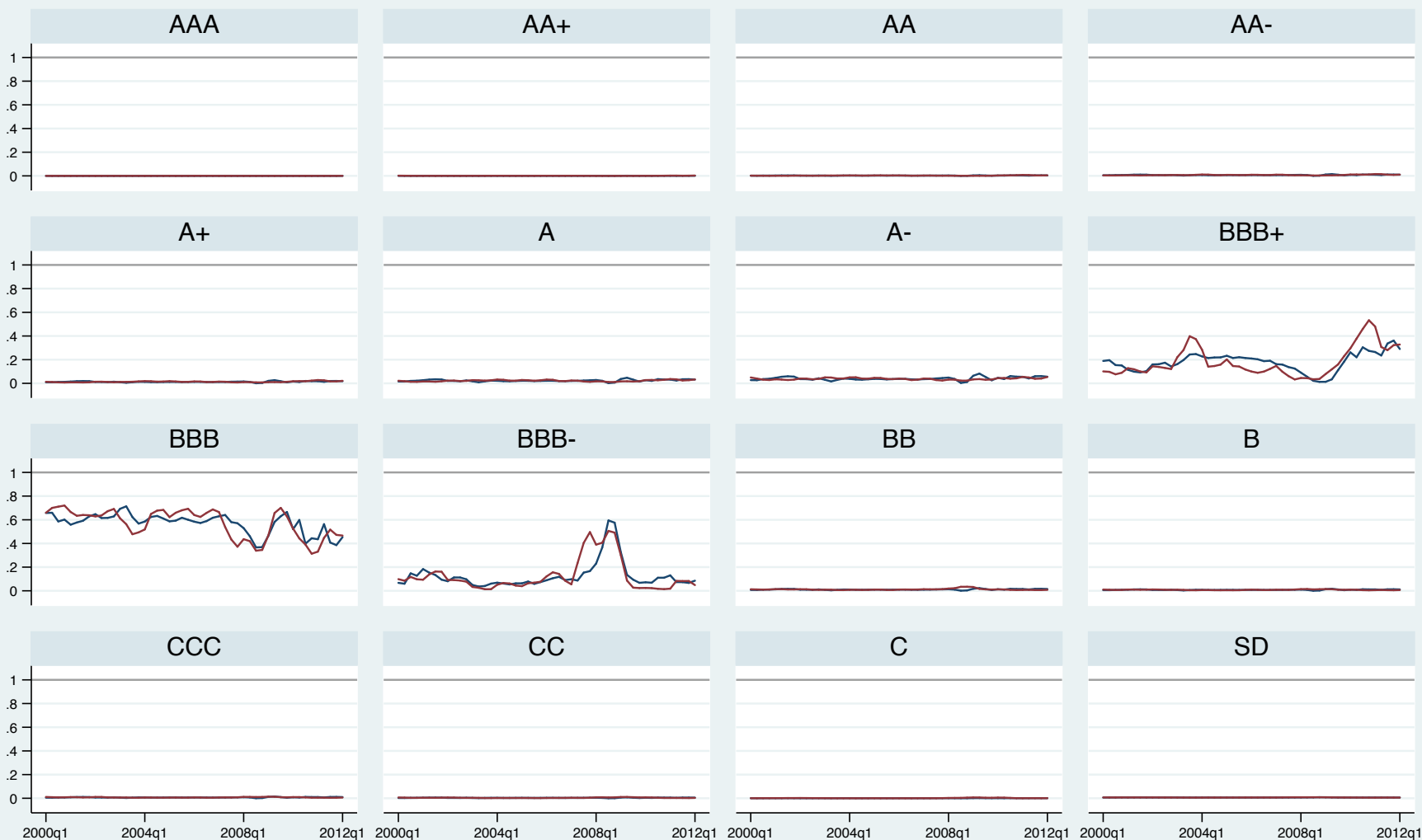
blue: predicted series, red: historical



Graphs by tgt

Predicted transitions from BBB

blue: predicted series, red: historical



Graphs by tgt

- The fractional logit approach to modeling the most important elements of the transition matrix is quite successful, and produces predictions in the probability metric without retransformation bias.
- As the diagonal, super- and sub-diagonals of the transition matrix capture a sizable fraction of transition probabilities, the problem of estimating a time-varying transition matrix is greatly simplified.
- This approach allows the introduction of relevant macroeconomic drivers which affect the transition probabilities directly, as well as influencing the transition coefficients from each AQ rating class.

- The fractional logit approach to modeling the most important elements of the transition matrix is quite successful, and produces predictions in the probability metric without retransformation bias.
- As the diagonal, super- and sub-diagonals of the transition matrix capture a sizable fraction of transition probabilities, the problem of estimating a time-varying transition matrix is greatly simplified.
- This approach allows the introduction of relevant macroeconomic drivers which affect the transition probabilities directly, as well as influencing the transition coefficients from each AQ rating class.

- The fractional logit approach to modeling the most important elements of the transition matrix is quite successful, and produces predictions in the probability metric without retransformation bias.
- As the diagonal, super- and sub-diagonals of the transition matrix capture a sizable fraction of transition probabilities, the problem of estimating a time-varying transition matrix is greatly simplified.
- This approach allows the introduction of relevant macroeconomic drivers which affect the transition probabilities directly, as well as influencing the transition coefficients from each AQ rating class.

A SUR approach to modeling defaults

- We now present an alternative empirical approach based on modeling the observed default rates for each level of asset quality in a portfolio of wholesale loans.
- This dynamic model expresses the default rate for asset quality i , $AQ_{i,t}$, in terms of lagged values of $AQ_{i,t}$, $AQ_{i-1,t}$ and $AQ_{i+1,t}$ in addition to relevant macroeconomic factors.
- The model is fit as a system of Seemingly Unrelated Regressions (SUR), a generalized least squares technique (Zellner, *JASA* 1962) that allows contemporaneous correlation of errors to be used to gain efficiency in the estimation: in Stata, the `sureg` estimation command.

A SUR approach to modeling defaults

- We now present an alternative empirical approach based on modeling the observed default rates for each level of asset quality in a portfolio of wholesale loans.
- This dynamic model expresses the default rate for asset quality i , $AQ_{i,t}$, in terms of lagged values of $AQ_{i,t}$, $AQ_{i-1,t}$ and $AQ_{i+1,t}$ in addition to relevant macroeconomic factors.
- The model is fit as a system of Seemingly Unrelated Regressions (SUR), a generalized least squares technique (Zellner, *JASA* 1962) that allows contemporaneous correlation of errors to be used to gain efficiency in the estimation: in Stata, the `sureg` estimation command.

A SUR approach to modeling defaults

- We now present an alternative empirical approach based on modeling the observed default rates for each level of asset quality in a portfolio of wholesale loans.
- This dynamic model expresses the default rate for asset quality i , $AQ_{i,t}$, in terms of lagged values of $AQ_{i,t}$, $AQ_{i-1,t}$ and $AQ_{i+1,t}$ in addition to relevant macroeconomic factors.
- The model is fit as a system of Seemingly Unrelated Regressions (SUR), a generalized least squares technique (Zellner, *JASA* 1962) that allows contemporaneous correlation of errors to be used to gain efficiency in the estimation: in Stata, the `sureg` estimation command.

- In the empirical application, we have quarterly data for 1997–2012 on default rates for 16 asset quality ratings from a wholesale portfolio of C&I loans, where low ratings are of higher quality (AAA, AA+, AA...), with AQ16 denoting default (SD).
- Historical default rates vary considerably across asset quality ratings as well as over time.

Table: Selected descriptive statistics, 1997–2012

Rating	mean	s.d.	min	max
AAA	.0018172	.0016586	0.0	.0086655
A+	.0048467	.0025159	.0007106	.0095858
BBB	.0121260	.0063674	.0009095	.0227618
CCC	.0304619	.0171742	.0017301	.0758738

- In the empirical application, we have quarterly data for 1997–2012 on default rates for 16 asset quality ratings from a wholesale portfolio of C&I loans, where low ratings are of higher quality (AAA, AA+, AA...), with AQ16 denoting default (SD).
- Historical default rates vary considerably across asset quality ratings as well as over time.

Table: Selected descriptive statistics, 1997–2012

Rating	mean	s.d.	min	max
AAA	.0018172	.0016586	0.0	.0086655
A+	.0048467	.0025159	.0007106	.0095858
BBB	.0121260	.0063674	.0009095	.0227618
CCC	.0304619	.0171742	.0017301	.0758738

- In the empirical application, we have quarterly data for 1997–2012 on default rates for 16 asset quality ratings from a wholesale portfolio of C&I loans, where low ratings are of higher quality (AAA, AA+, AA...), with AQ16 denoting default (SD).
- Historical default rates vary considerably across asset quality ratings as well as over time.

Table: Selected descriptive statistics, 1997–2012

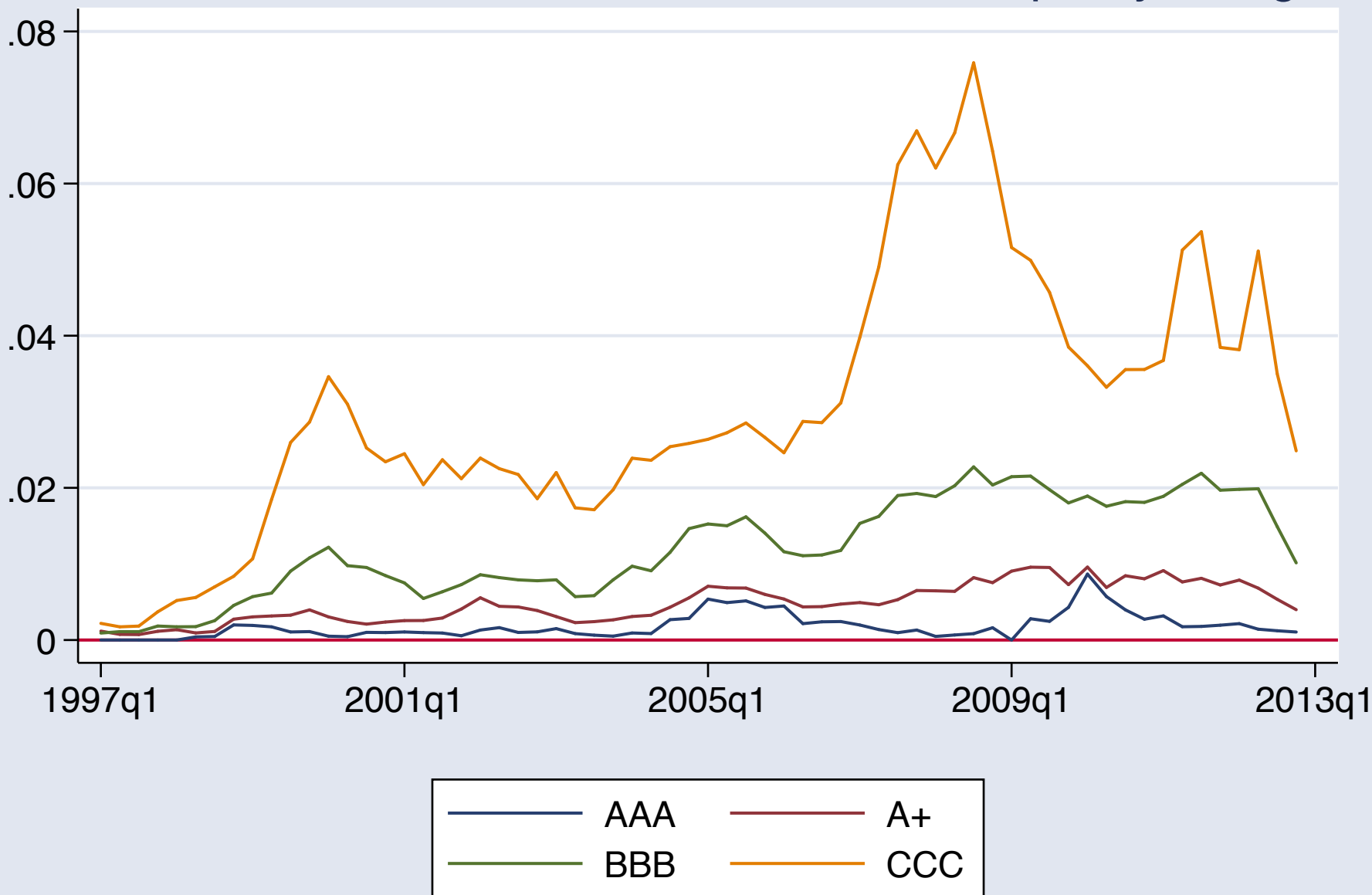
Rating	mean	s.d.	min	max
AAA	.0018172	.0016586	0.0	.0086655
A+	.0048467	.0025159	.0007106	.0095858
BBB	.0121260	.0063674	.0009095	.0227618
CCC	.0304619	.0171742	.0017301	.0758738

- In the empirical application, we have quarterly data for 1997–2012 on default rates for 16 asset quality ratings from a wholesale portfolio of C&I loans, where low ratings are of higher quality (AAA, AA+, AA...), with AQ16 denoting default (SD).
- Historical default rates vary considerably across asset quality ratings as well as over time.

Table: Selected descriptive statistics, 1997–2012

Rating	mean	s.d.	min	max
AAA	.0018172	.0016586	0.0	.0086655
A+	.0048467	.0025159	.0007106	.0095858
BBB	.0121260	.0063674	.0009095	.0227618
CCC	.0304619	.0171742	.0017301	.0758738

Historical default rates for selected asset quality ratings



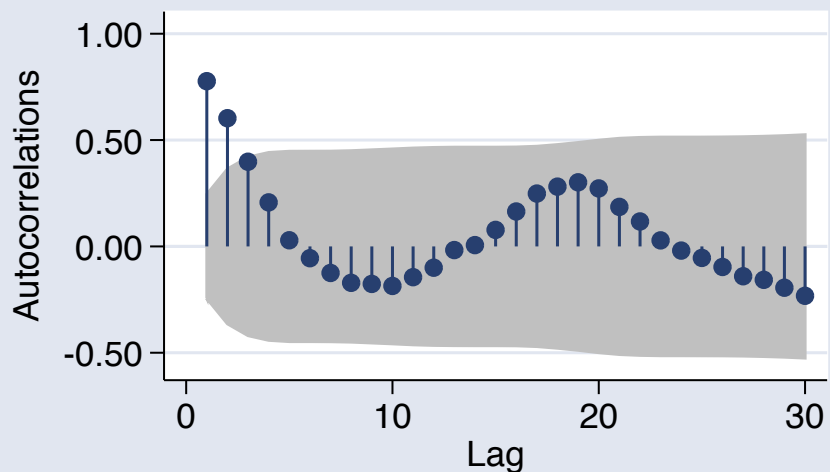
- The dynamic model we have implemented depends on the persistence of default rates for each rating class.
- We can examine that persistence by computing the autocorrelation functions for each default rate.
- The autocorrelation functions for selected AQ ratings show that between two and four quarterly lags of those ratings have autocorrelation coefficients significantly differing from zero.
- These findings suggest that several lags should be included in the SUR estimation to capture the persistence of default rates.

- The dynamic model we have implemented depends on the persistence of default rates for each rating class.
- We can examine that persistence by computing the autocorrelation functions for each default rate.
- The autocorrelation functions for selected AQ ratings show that between two and four quarterly lags of those ratings have autocorrelation coefficients significantly differing from zero.
- These findings suggest that several lags should be included in the SUR estimation to capture the persistence of default rates.

- The dynamic model we have implemented depends on the persistence of default rates for each rating class.
- We can examine that persistence by computing the autocorrelation functions for each default rate.
- The autocorrelation functions for selected AQ ratings show that between two and four quarterly lags of those ratings have autocorrelation coefficients significantly differing from zero.
- These findings suggest that several lags should be included in the SUR estimation to capture the persistence of default rates.

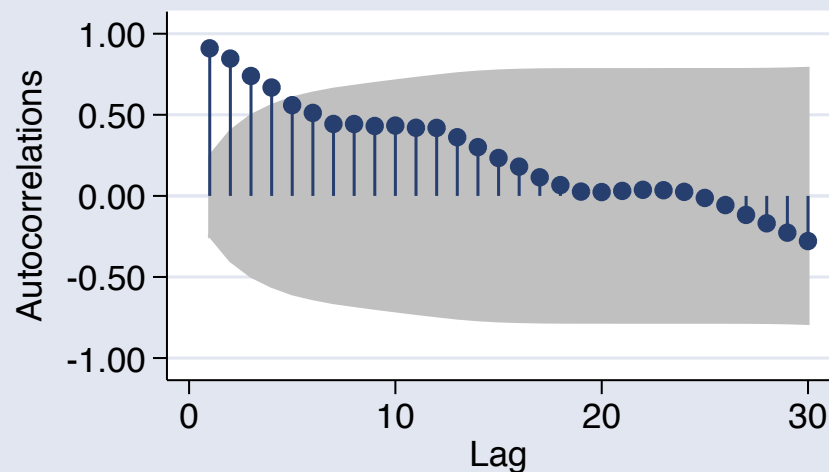
- The dynamic model we have implemented depends on the persistence of default rates for each rating class.
- We can examine that persistence by computing the autocorrelation functions for each default rate.
- The autocorrelation functions for selected AQ ratings show that between two and four quarterly lags of those ratings have autocorrelation coefficients significantly differing from zero.
- These findings suggest that several lags should be included in the SUR estimation to capture the persistence of default rates.

AAA



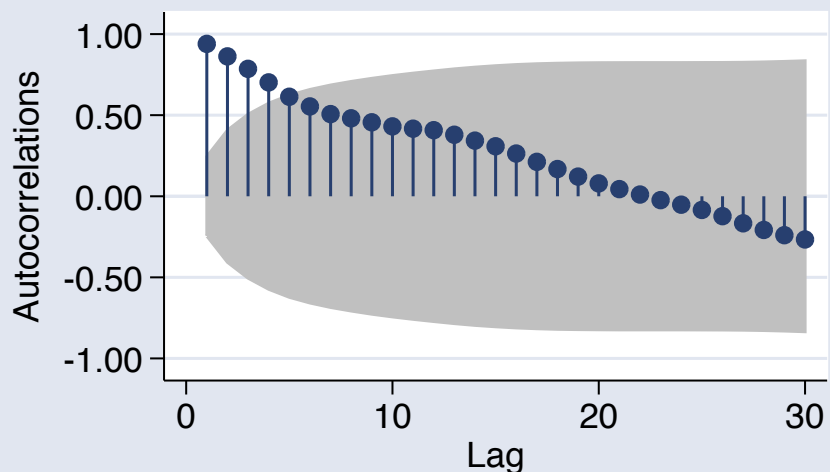
Bartlett's formula for MA(q) 95% confidence bands

A+



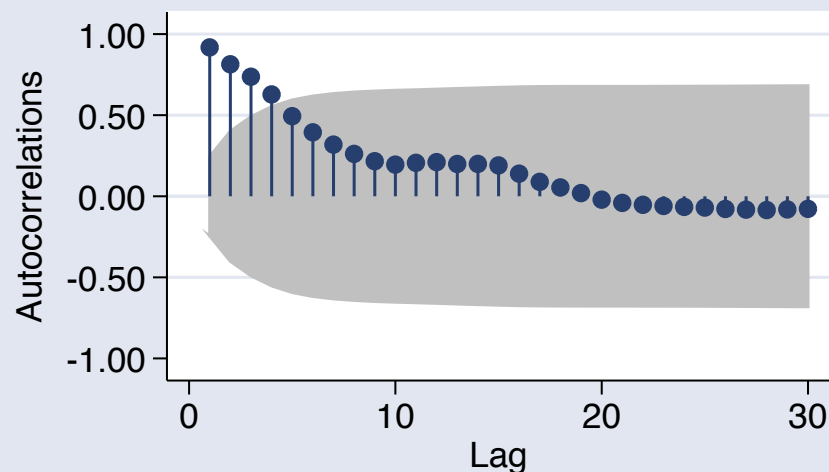
Bartlett's formula for MA(q) 95% confidence bands

BBB



Bartlett's formula for MA(q) 95% confidence bands

CCC



Bartlett's formula for MA(q) 95% confidence bands

To capture common factors influencing all borrowers, we include a set of macroeconomic factors in each equation:

- Real GDP growth, YoY, three-quarter centered moving average
- Unemployment \times the change in unemployment from t to $t + 1$
- Average weekly hours in manufacturing

To capture common factors influencing all borrowers, we include a set of macroeconomic factors in each equation:

- Real GDP growth, YoY, three-quarter centered moving average
- Unemployment \times the change in unemployment from t to $t + 1$
- Average weekly hours in manufacturing

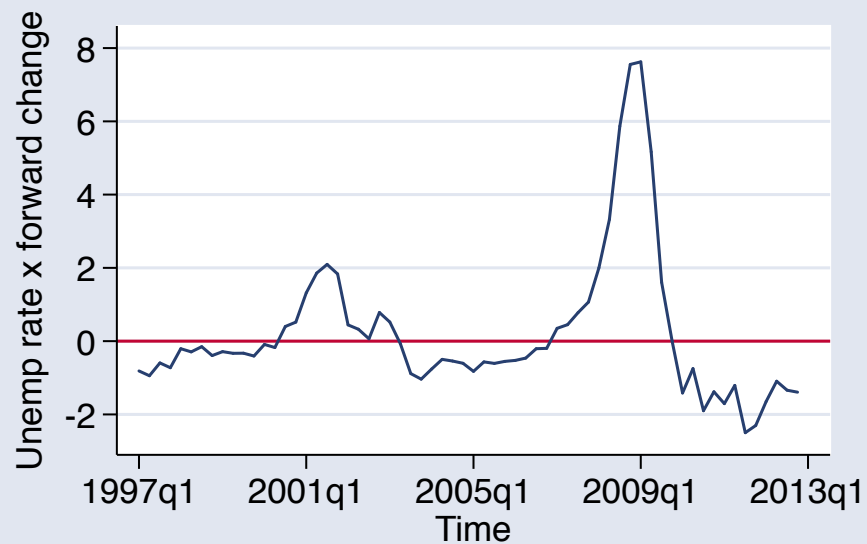
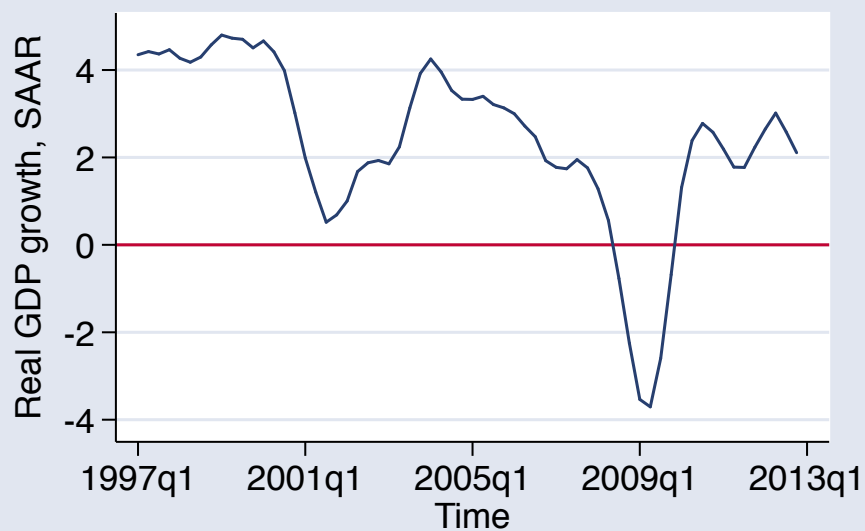
To capture common factors influencing all borrowers, we include a set of macroeconomic factors in each equation:

- Real GDP growth, YoY, three-quarter centered moving average
- Unemployment \times the change in unemployment from t to $t + 1$
- Average weekly hours in manufacturing

To capture common factors influencing all borrowers, we include a set of macroeconomic factors in each equation:

- Real GDP growth, YoY, three-quarter centered moving average
- Unemployment \times the change in unemployment from t to $t + 1$
- Average weekly hours in manufacturing

SUR macro variables



- The explanatory power of the estimated model is very good, with individual equations' R^2 values of 0.68 for AAA-rated borrowers, between 0.79 and 0.88 for AA+ through A rated borrowers, and generally above 0.92 for the remaining asset quality ratings.
- Tests for exclusion of the macroeconomic factors reject their null hypothesis at the 90% or 95% level.
- The SUR technique improves upon equation-by-equation OLS when the residual correlation matrix contains sizable off-diagonal elements.
- The Breusch–Pagan test for a diagonal correlation matrix rejects its null hypothesis at any level of confidence, implying that there are important contemporaneous correlations among the equations' errors: common shocks not captured by the macro factors.

- The explanatory power of the estimated model is very good, with individual equations' R^2 values of 0.68 for AAA-rated borrowers, between 0.79 and 0.88 for AA+ through A rated borrowers, and generally above 0.92 for the remaining asset quality ratings.
- Tests for exclusion of the macroeconomic factors reject their null hypothesis at the 90% or 95% level.
- The SUR technique improves upon equation-by-equation OLS when the residual correlation matrix contains sizable off-diagonal elements.
- The Breusch–Pagan test for a diagonal correlation matrix rejects its null hypothesis at any level of confidence, implying that there are important contemporaneous correlations among the equations' errors: common shocks not captured by the macro factors.

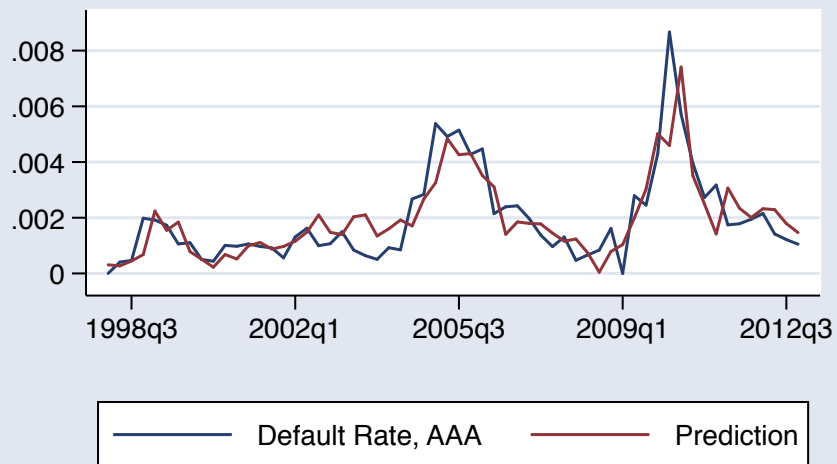
- The explanatory power of the estimated model is very good, with individual equations' R^2 values of 0.68 for AAA-rated borrowers, between 0.79 and 0.88 for AA+ through A rated borrowers, and generally above 0.92 for the remaining asset quality ratings.
- Tests for exclusion of the macroeconomic factors reject their null hypothesis at the 90% or 95% level.
- The SUR technique improves upon equation-by-equation OLS when the residual correlation matrix contains sizable off-diagonal elements.
- The Breusch–Pagan test for a diagonal correlation matrix rejects its null hypothesis at any level of confidence, implying that there are important contemporaneous correlations among the equations' errors: common shocks not captured by the macro factors.

- The explanatory power of the estimated model is very good, with individual equations' R^2 values of 0.68 for AAA-rated borrowers, between 0.79 and 0.88 for AA+ through A rated borrowers, and generally above 0.92 for the remaining asset quality ratings.
- Tests for exclusion of the macroeconomic factors reject their null hypothesis at the 90% or 95% level.
- The SUR technique improves upon equation-by-equation OLS when the residual correlation matrix contains sizable off-diagonal elements.
- The Breusch–Pagan test for a diagonal correlation matrix rejects its null hypothesis at any level of confidence, implying that there are important contemporaneous correlations among the equations' errors: common shocks not captured by the macro factors.

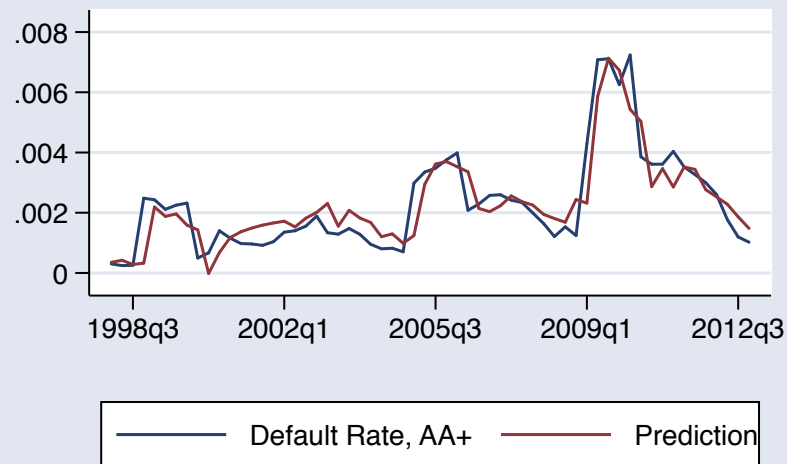
In-sample comparisons of the actual and predicted default rates for each asset quality rating show that the model is able to capture the trajectories of these series quite well through the period of the financial crisis.

Modeled Default Rate

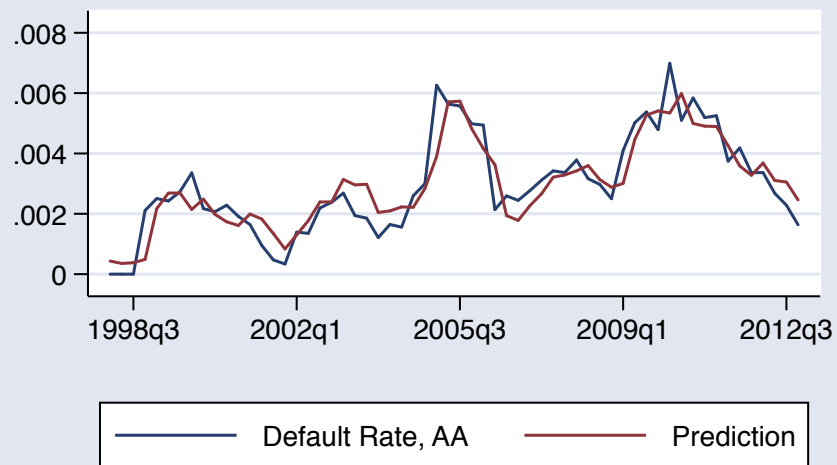
AAA



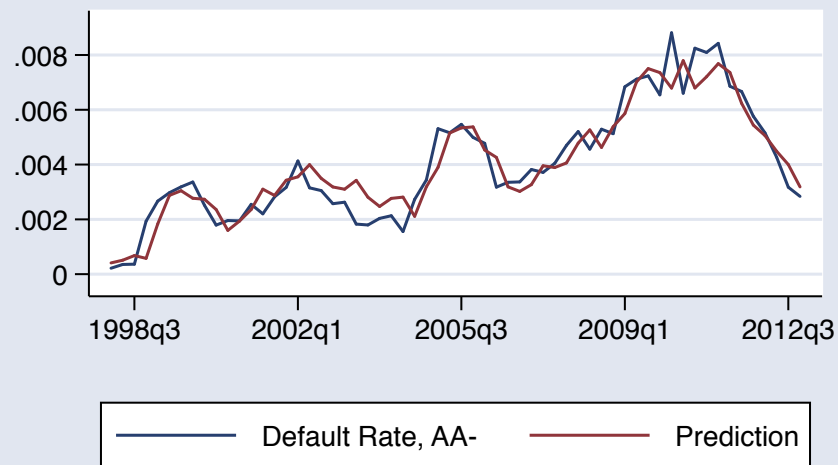
AA+



AA

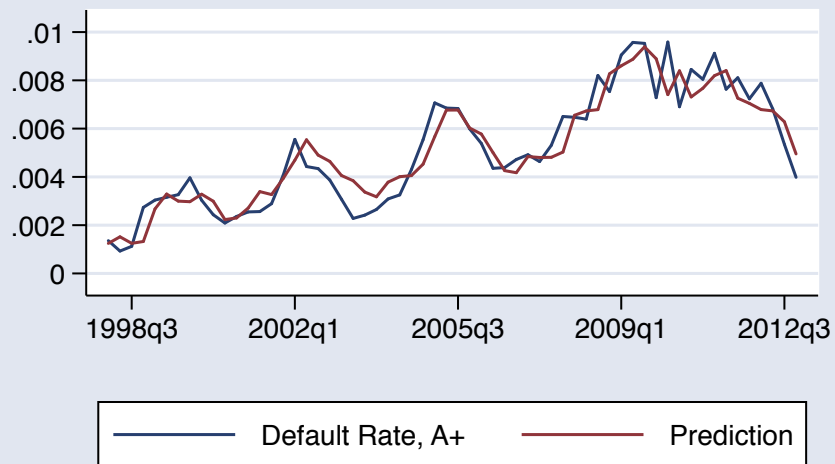


AA-

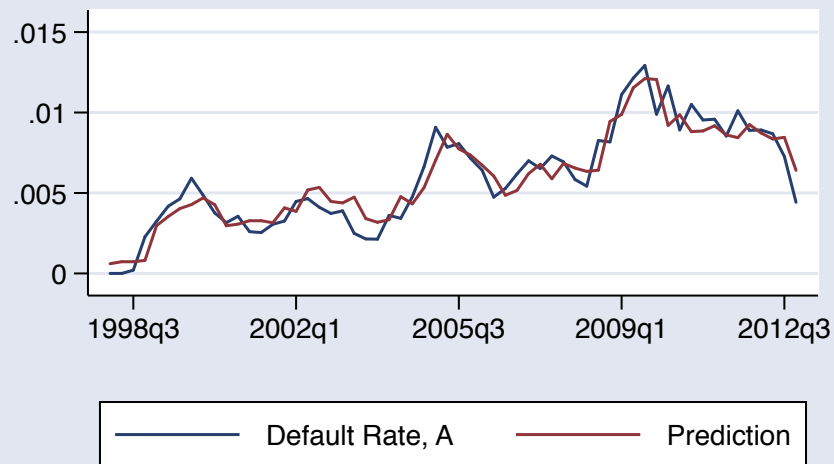


Modeled Default Rate

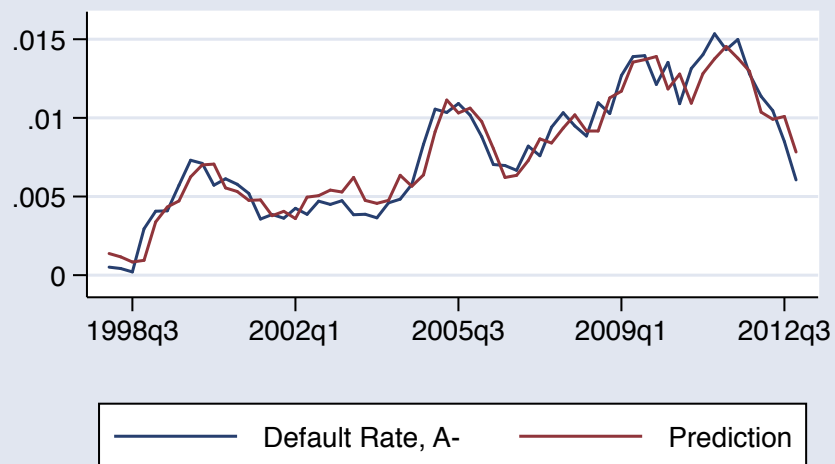
A+



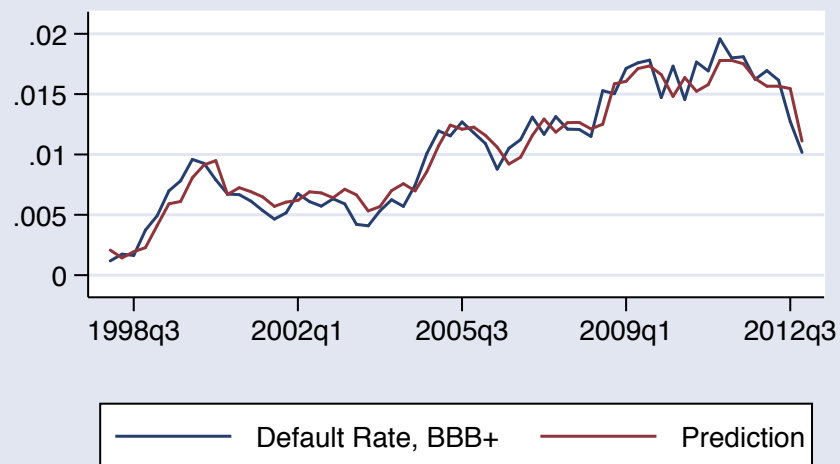
A



A-

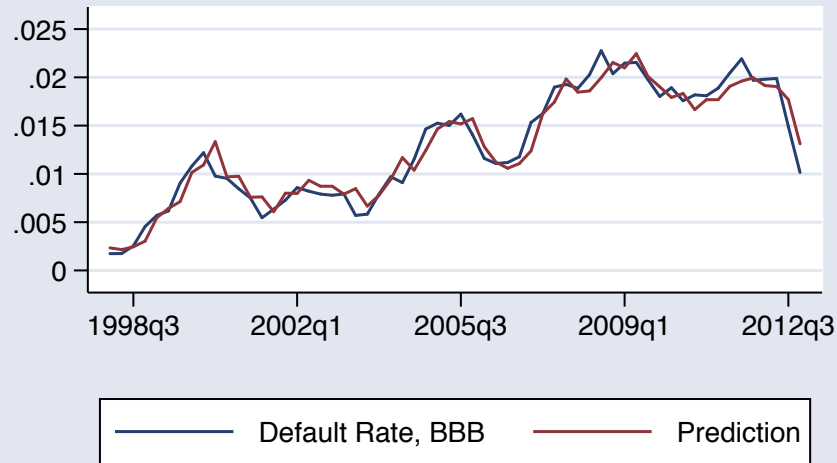


BBB+

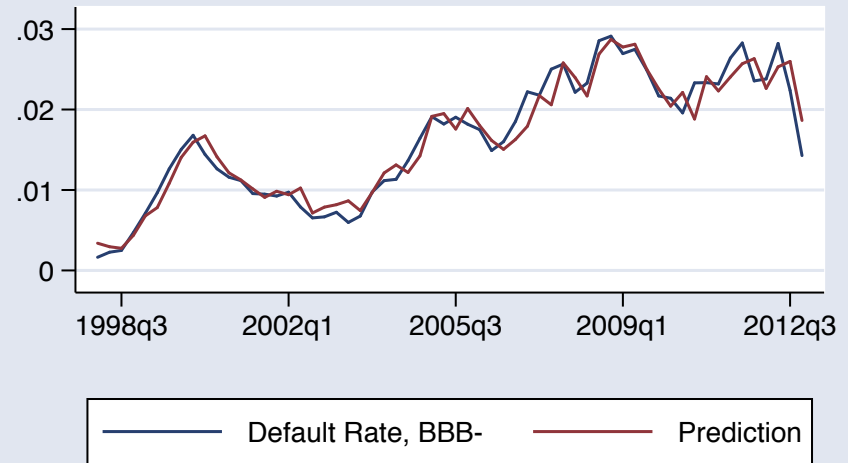


Modeled Default Rate

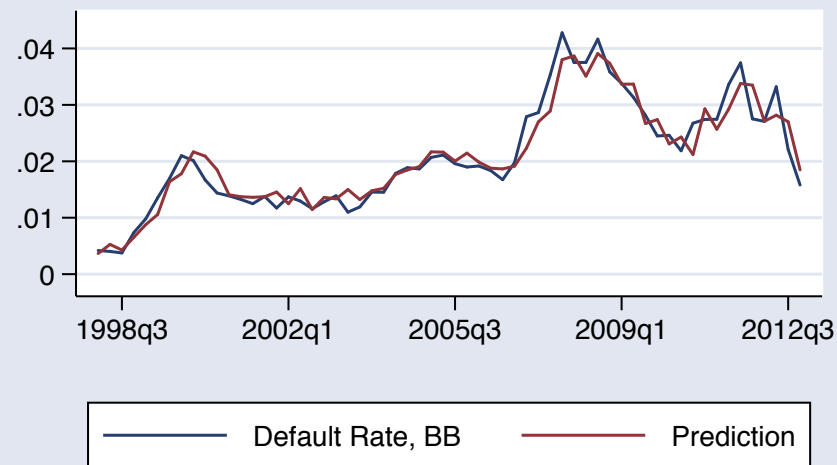
BBB



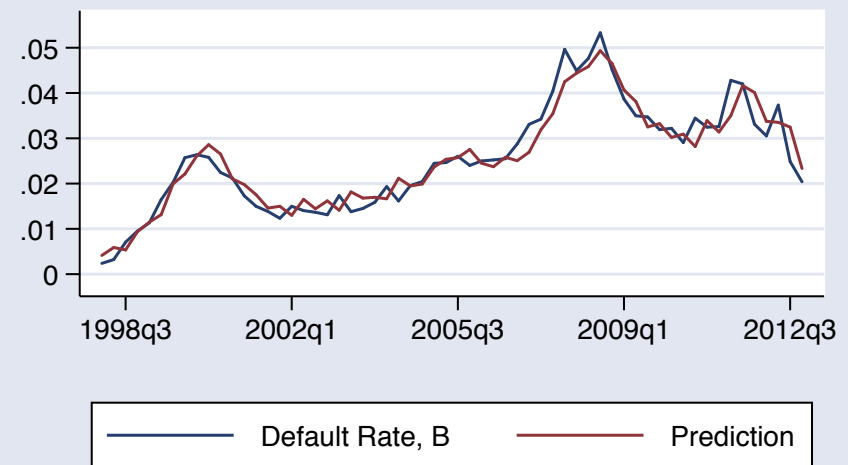
BBB-



BB

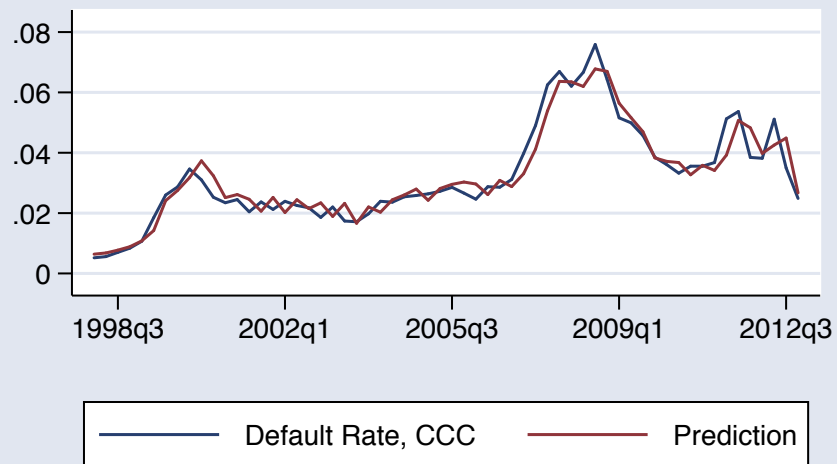


B

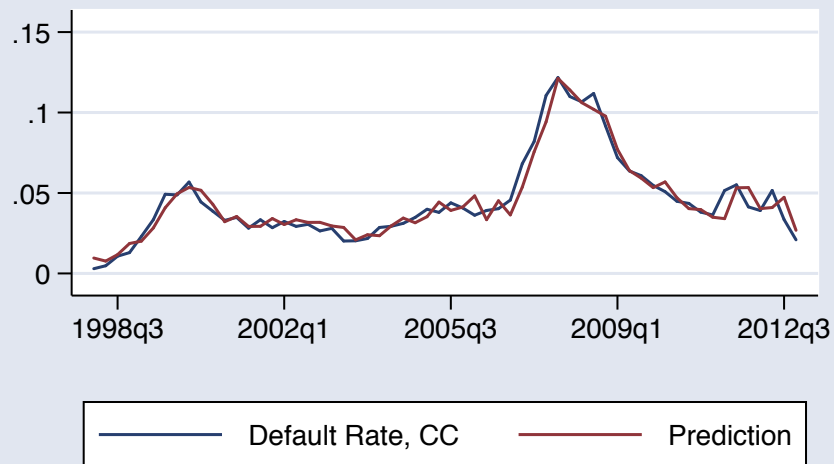


Modeled Default Rate

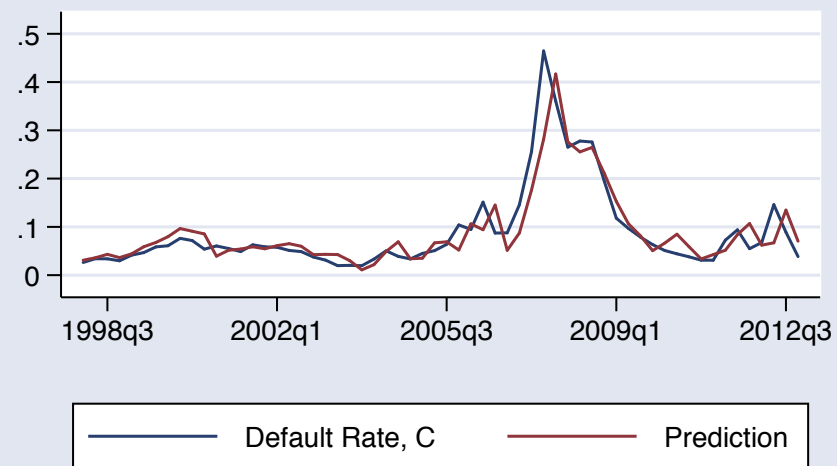
CCC



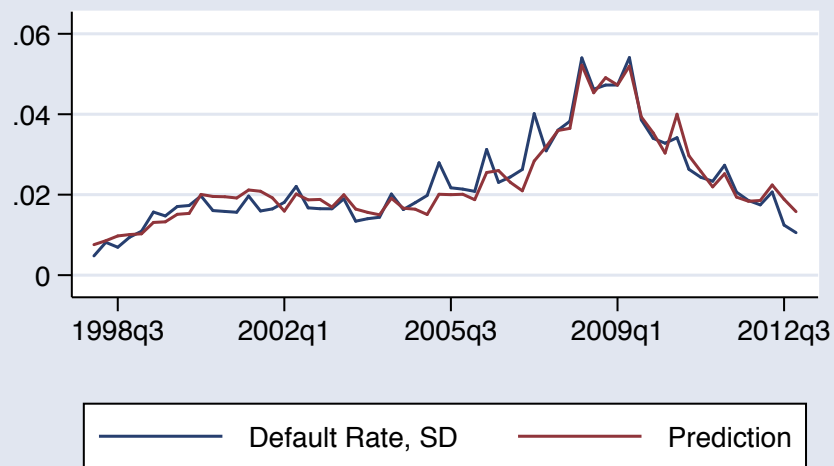
CC



C



SD



- The SUR model jointly captures the movements of default rates across asset quality ratings by exploiting the persistence in default rates for a given asset quality and its neighbors.
- The model can be augmented with macroeconomic factors to capture common shocks affecting all borrowers.
- The SUR technique makes use of the contemporaneous correlation across equations' errors to gain efficiency in estimation relative to equation-by-equation OLS.

- The SUR model jointly captures the movements of default rates across asset quality ratings by exploiting the persistence in default rates for a given asset quality and its neighbors.
- The model can be augmented with macroeconomic factors to capture common shocks affecting all borrowers.
- The SUR technique makes use of the contemporaneous correlation across equations' errors to gain efficiency in estimation relative to equation-by-equation OLS.

- The SUR model jointly captures the movements of default rates across asset quality ratings by exploiting the persistence in default rates for a given asset quality and its neighbors.
- The model can be augmented with macroeconomic factors to capture common shocks affecting all borrowers.
- The SUR technique makes use of the contemporaneous correlation across equations' errors to gain efficiency in estimation relative to equation-by-equation OLS.

The bottom line

In conclusion, Stata has become an essential element of the toolkit at CFG's Risk Analytics group. Although other statistical tools and databases are also used for data management tasks, the group's econometric modeling is heavily Stata-oriented due to the program's capabilities, programmability, cost-effectiveness and overall ease of use.