Credit risk migration rates modelling as open systems II: A Simulation Model and IFRS9-baseline principles

DOI:10.1016/j.strueco.2019.06.013

Credit risk migration rates modelling as open systems II: A Simulation Model and IFRS9-baseline principles

S. Landini[a,*], M. Uberti[b], S. Casellina[c]

[a] *IRES Piemonte, Socioeconomic Research Institute of Piedmont, via Nizza 18, 10125, Turin, Italy.*
[b] *Department of Management: Università degli Studi di Torino, Turin, Italy.*
[c] *Banca d'Italia: Regulation and Supervisory Policies Department, Rome, Italy.*

Abstract

In 2014 the *International Accounting Standards Board* (IASB) promulgated the current *International Financial Reporting Standards 9 - Financial Instruments* (IFRS9) that draw new lines for an *ex-ante*, reliable, unified and well-balanced credit risk assessment. Among others, two principles are of interest to this paper: that of *segmented* and *prospective* estimation of *expected credit losses*. Within the frame of a micro-simulation approach, this paper focuses on these issues while considering the evolution of a bank portfolio. The paper presents an algorithmic procedure developed on a realistic dynamic credit risk migration rates modelling of a portfolio as an open system with entries and exits that is consistent with the *segmented* and *prospective* IFRS9 principles. Although operating at the aggregate level of the migration matrix, combining accounting principles inspired to those of the IFRS9-baseline with the open systems modelling, the main conclusion is that it allows for a more reliable provision and *ex-ante* and forward-looking estimation of expected losses.

Keywords: Credit risk; Migration rates models; Micro-simulation; Expected loss; Accounting standards.
JEL: C15, C18, C53, C63, G11, G13, G17, G18, G21, G24, G28, G31, G34, G38, E47.

1. Introduction and motivation

The financial crisis of 2007 made clear to regulators that the practices of accounting losses and provision against the implied risk were too weak as they did not timely recognize deterioration in creditworthiness: the Incurred Loss criterion, that accounts for losses due to deteriorations that have already happened, turnes out not being applicable to estimate future reserves before a so-called *trigger-event* happens. On April 2nd 2009

*Corresponding author
Email addresses: landini@ires.piemonte.it (S. Landini), mariacristina.uberti@unito.it (M. Uberti), simoneenrico.casellina@bancaditalia.it (S. Casellina)

the G20 published *The Declaration on Strengthening the Financial System* [1] that stimulated regulators at providing new standards, either on the accounting and on the prudential side. In 2014 the *International Accounting Standards Board* (IASB) promulgated the current *International Financial Reporting Standards 9 - Financial Instruments* (IFRS9).

Since January 1st 2018 the *IFRS9-Financial Instruments* (see [2], [3], [4],[5], [6]), issued in 2014 by the *IAS B*, substitutes the previous *IASB39-Financial Instruments*. The IFRS9 introduces new challenges for a more reliable *ex-ante* and well-balanced assessment of banking book credit risk, that has to be *segmented* and *forward-looking* at estimating bank's potential losses. Since the delayed recognition of credit losses on loans was identified as a weakness in existing accounting standards, the IASB has introduced an Expected Loss (*EL*) model that requires more timely recognition of credit losses.

According to the accounting standard (IFRS9), the *ex-ante* assessment concerns expected losses due to impairment events of debtors, each of which must be appropriately classified to measure risk. Also, the estimate of potential losses should not be based only on past and contingent events (Incurred Losses) but, most of all, according to a forward-looking approach based on future macroeconomic scenarios. The *ex-ante* assessment should be *segmented* to the extent that, based on the risk related to the creditworthiness, each position in the banking portfolio should be classified into appropriated *buckets*, each specific to a recognised business model and the time-regime at which expected defaults may happen in order to assess a reliable economic capital estimate, not depending on the contractual maturity of the instrument. In this framework, it is necessary a re-thinking of the theoretical approach that involves migration rates matrices in credit risk modelling as the *standard approach* (among others see [7, 8] and literature cited therein, see also [9] and [10]) is based on a set of items: (a) closed-sample cohorts; (b) *a-priori* absorbing *default* grade; (c) time-average migration rates matrix; (d) Makovian hypothesis; (e) eigen-decomposition of the migration matrix to forecast future trends.

In [11] an original methodological approach (ESL model) has been recently proposed that is innovative with respect to the *standard method* of migration rates modelling for credit risk from the point of view of a bank where, usually, lenders are assumed to rate borrowers according to a closed-sample approach that, *a priori*, assumes the default state D as absorbing. As widely known and experienced by analysts, the *standard method* could not predict other than the whole portfolio will collapse to default even in the short-medium run, so predicting an overestimation of the economic capital provision that, in the end, would induce the collapse of the considered credit-line, if not of the bank too. As it can be understood, with the standard approach the greater the portfolio the more the bankruptcy of the lender is likely expected in the short run.

Clearly, this cannot meet the current normative requirements on the economic capital provision, as a matter of fact, it is a puzzling point for analysts, banks and supervisors.

To overcome these limitations, also *at-default* debtors coming back *in-bonis* (i.e. *cure events*) are taken into account while the standard practice considers the default state as absorbing. Moreover, the non-standard approach developed by [11] is based on assumptions motivated by empirical observation and theoretical outcomes and it considers a portfolio as a sample allowing for entries (E), migrations of stayers (S) and exits (L) as well: i.e. a dynamically changing open system with renewal.

A micro-simulation based approach of the credit portfolio of a bank is then developed as an open system with renewal and migrations, suggesting three main conclusions in terms of economic capital provision: (a) based on the Markovian hypothesis with *a-priori* absorbing state *at-default*, the standard closed-sample approach has to be abandoned not to predict lenders' bankruptcy by construction; (b) to meet more reliable estimates along with the new regulatory standards, the sample to estimate migration matrices has to be opened either to entries and exits; (c) the usual static eigen-decomposition procedure to forecast migration rates has to be replaced with more reliable practices, for which this paper provides a first attempt to meet practical needs.

The ELS model developed in [11] simulates a portfolio of identical financial instruments with renewal: open to entries and exits beyond the standard stays. Besides producing macro and meso-data, the simulation model also produces micro-data in historical series, holding memory of the past evolution of every contract in the portfolio, from the subscription to the expiry date. Therefore, the micro-simulation framework of the ELS model can be considered as the DGP (data generating process) of a portfolio interpreted as a system open to entries and exits. Moreover, the ESL model describes a portfolio as an open system that dynamically renews through time along with the macroeconomic state of the whole economy. Therefore, the characteristic of the ESL model make it a suitable device to develop, at aggregate level, a modelling that may be a starting point toward the above mentioned IFS9-baseline principles.

For simplicity, the ESL model is based on a *simplified and augmented master-scale*[1] of rating classes or states: $\Lambda_0 = \{\lambda_0 = E, \lambda_1 = A, \lambda_2 = B, \lambda_3 = C, \lambda_4 = D, \lambda_5 = L\}$, where E identifies the entries, A is the best creditworthiness class, B is the medium-high, C is the medium-low, D is the default class and L identifies the exits. Therefore, the standard grades of creditworthiness are actually given by a master-scale $\Lambda = \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ which is augmented with two classifications obtaining $\Lambda_0 = E \cup \Lambda \cup L$ in order to handle

[1] *Simplified* as just few standard grades are considered, *augmented* as it includes entries (E) and exits (L).

the entries and the exits of the portfolio[2].

By applying the ESL model, at least at the baseline level and to the extent of simulation purposes at aggregate level, the aim of this paper is to specify a mathematical procedure that takes into account two of the main principles introduced by IFRS9, namely that of segmentation and prospectiveness. More precisely, as the IFRS9 may be structured in three pillars (classification and measurement, impairment, hedge accounting), we focus our modelling just on the impairment pillar, that involves *expected credit losses* for positions that are grouped into three buckets, each characterized by specific rules: technical and normative details on the implementation of the standard can be found in [2, 3, 4, 5, 6]. Therefore, after [11], in this paper the ESL modelling approach is combined with baseline principles inspired to those of the IFRS9 as regarding the *segmented* and *prospective* estimates of expected losses to evaluate their predictive capacity. From a practical point of view, this paper may also provide a first methodological approach to comprehend some of the IFRS9 issues in terms of modelling and simulation methods. The main obtained results show that combining the ESL modelling with a simplified set of principles inspired to the IFRS9-baseline can provide a more effective estimation of expected credit losses as it overcomes some of the limitations of the *standard approach*, so providing a more reliable capital provision. Of course, this is not a definitive result but a starting point for further research.

The remainder of the paper is organized as follows. The theoretical model assumptions are introduced in Section 2. In Section 3 the micro-simulation modelling procedures are set up: the estimation procedure is defined in Section 3.1, the identification one is described in Section 3.2, the extrapolation one is summarized in Section 3.3 and the results are collected in Section 3.4. In Section 4 the extrapolation of the migration probability matrices is studied where, in particular, sections 4.1 and 4.2 are devoted to the entry and the exit migration rates by grades and to the cure and the default migration rates by grades, respectively. In Section 5 absolute value migration matrices are extrapolated to meet the accounting standard requirements conditionally to a future scenario with a procedure made of three phases: Phase 1: the whole (Section 5.1); Phase 2: the parts (Section 5.2) and Phase 3: the cells (Section 5.3). In Section 6 simplified baseline principles inspired to those of the IFRS9 are introduced and future dynamics of the portfolio is extrapolated consistently with a segmentation in three buckets that, at an aggregate level, approximate the rationale of the IFRS9, and the expected loss of the portfolio is estimated accordingly. Section 7 concludes this note. Figure 1 in Section 3 summarises the whole procedure while providing a graphical abstract of the paper ~~that may~~

[2]With respect to the Assumption 9 and the equation (25) in [11], a further grade C is here introduced.

~~help the reader~~.

2. Theoretical model hypotheses

This section introduces the main assumptions of the simulation methodology involved as well as those of the theoretical modelling developed by [11].

A time schedule $\mathbb{T}_{SIM} = \{1, \dots, t_0, \dots, T, \dots, T^*\} \subset \mathbb{N}$ is considered where three time windows are identified:

- *training window* $\mathbb{T}_0 = \{1, \dots, t_0\} \subset \mathbb{T}_{SIM}$: it indexes the sequence of events with which the simulation model is initialized; these data will be neglected in the next phases;
- *observation window* $\mathbb{T}_{OBS} = \{t_0 + 1, \dots, T\} \subset \mathbb{T}_{SIM}$: during this time interval, the actual simulation takes place using the ESL model that will be considered as the historical database for the evolution of the bank's portfolio, but also for the macro variables that have cyclically accompanied it; the observation window coincides with the *estimation* $\mathbb{T}_{EST}$ and *identification* $\mathbb{T}_{IDF}$ windows;
- *extrapolation window* $\mathbb{T}_{EXT} = \{T + 1, \dots, T^*\} \subset \mathbb{T}_{SIM}$: the simulation proceeds also during this period mainly for control purposes but, in estimating the necessary quantities, data simulated in this part will not be taken into account to be instead extrapolated as if they were unknown.

The overall simulation covers the whole time schedule $\mathbb{T}_{SIM} = \mathbb{T}_0 \cup \mathbb{T}_{OBS} \cup \mathbb{T}_{EXT}$.

For the ease of comprehension, it is assumed that each date represents an annual yield but, more generally and properly, it can as well be assumed that each point in time refers to a *reporting-date*, consistently with the current IFRS9 accounting standard. Accordingly: by setting $t_0 = 10$ the initialization phase covers 10 years so that $t_0+1 = 11$ is the year since portfolio of a single credit line is assumed to be observed consistently with assumptions 1 and 2 in [11]: i.e., as a simplification we consider a single type of financial instrument and all the contracts are assumed to be signed with the same maturity. By setting $T = 40$ the historical simulation concerns 40 years and the last year is considered as the *present*, finally setting $T^* = 50$ it is assumed to have a *future* decade on which to make forecasts extrapolating the dynamics of aggregate quantities involved in the model. It is worth pointing out that, currently, no such long time series are available as regarding migration rates matrices and portfolios; also, we tested the model with different and longer periods without detecting instabilities. The here proposed yields are intended for exhibition purposes only, as well as the whole simulation procedures deal with artificial data and results.

Throughout time schedule $\mathbb{T}_{SIM}$, the ESL model simulates the following quantities:

- Macro level:[3] the GDP $Y_t \geq 0$, the related growth rate $\gamma_t \in \mathbb{R}$ and a latent systematic factor $Z_t \in \mathbb{R}$[4];
- Micro level:[5] individual creditworthiness $W_{i,t} \geq 0$ and idiosyncratic shocks $Z_{i,t}$, the portfolio $\mathbb{F}_t = \mathbb{E}_t \cup \mathbb{S}_t = \mathbb{E}_t \cup \mathbb{F}_{t-1} \setminus \mathbb{L}_t$ where $\mathbb{E}_t$ is the sample of new entries, $\mathbb{S}_t$ is that of the stayers and $\mathbb{L}_t$ is that of the leavers;
- Meso level:[6] the migration matrices in absolute value $\mathbf{N}_t = \{N_{h,k,t}\}$, $N_{h,k,t} \in \mathbb{N}_0$, and probability of migration $\mathbf{P}_t$.

Within the simulation framework, the observation phase coincides with the *in-sample* estimation phase of the systematic factor Z_t that will be introduced in a while. That is, during $\mathbb{T}_{EST} \equiv \mathbb{T}_{OBS}$ the Z_t^{EST} series will be obtained: being hidden, this factor is not known nor knowable while, at the same time, somehow axiomatically it is considered to exist, so the problem of how to estimate it will be proposed and solved consistently to simulated data that, provided their availability, may be also substituted with real data. Moreover, the *estimation* phase coincides with the *identification* phase of a simple stochastic equation to interpret the estimated systematic factor so that it can be subsequently projected with *extrapolation* to the future on the basis of data that can be assumed to be made available by some institution, i.e. during $\mathbb{T}_{IDF} \equiv \mathbb{T}_{EST}$ there will be the identification an econometric model for the $Z_t^{IDF} \in \mathbb{R}$ series. Finally, after the observation-estimate-identification, the $\mathbb{T}_{EXT}$ extrapolation phase will follow for the *expected dynamics* of Z_t^{EXT}, $\mathbf{N}_t^{EXT}$, $\mathbf{P}_t^{EXT}$ conditionally to future scenario forecasts that could be produced by some institution.
It is worth mentioning that, in the ESL model, the sample $\mathbb{E}_t$ of entries as well as the individual creditworthiness $W_{i,t}$ are influenced[7] by the cycle of Y_t and the related dynamics γ_t, which are simulated but nothing forbids to use real data on the GDP and the related rate of growth in place of the simulated ones[8].

---

[3] In [11] see eq. (9), (10) and (12), where in the latter Z_t^1 has here been written as Z_t for ease of notation convenience.

[4] See [12], [11] and [8]; among others, [13] refers to this as a systematic risk-factor.

[5] In [11] see eq. (14), (13) and (7) respectively.

[6] In [11] see tables 6 and 7 respectively.

[7] In [11] see eq. (6) and (11) for entries and also eq. (12) and (14) for creditworthiness.

[8] Noticeably, IFRS9 does not limit the modelling to the GDP or to a single quantity. More sophisticated models can be specified to include other macroeconomic parameters like the unemployment and the inflation rate. Also, the modelling can be developed at different scales, e.g. national, regional or local. Not only, depending on the specific portfolio, different indices about the economic performance of specific sectors of activity can be involved, especially with the aim of considering the so-called *concentration* effects. In this paper we maintained the modelling at a very simple level of specification.

In this context, we decided to proceed within a completely simulative framework because, for the purpose of better consistency, to use real data on macro quantities one should also have access to a series of real migration matrices with entries and exits, i.e. to the micro-data of the creditworthiness of a portfolio. This is not possible for at least two reasons: not for many years, and only for a few large institutions, migration matrices have been recently prepared and micro-data have been stored in a structured and organized way for the purposes of recent accounting standards (for a recent review about the *state of art* see [3]) and, even if the collections were more extensive and the series were longer, the detailed migration matrices and the micro-data that generated them would, however, be accessible only to banks, which own them, and to the supervisory authorities. Nevertheless, although within the frame of a simplified simulation context, the methodology proposed in this paper constitutes a methodological basis for the purposes of an *ex-ante* credit risk assessment according to accounting and prudential principles.

In this regard, it is finally necessary to confirm an important assumption of the ESL model. Since the expected loss (EL) is defined as a function of (a) the exposure at default (EAD), (b) the loss given default (LGD) and (c) the probability of default (PD), i.e. $EL = f(EAD, LGD, PD)$, and because there is not yet sufficient agreement on the method of calculation of EAD and LGD, especially for hybrid portfolios, Assumption 1 in [11] is maintained. That is, we assume the portfolio is composed of contracts signed for a single type of financial instrument, which allows considering EAD and LGD as constants that can be neglected. Therefore, assuming a homogeneous portfolio implies a strong simplification: in fact, EL will depend only on PD, i.e. $EL = f(PD)$, an information that we can process in the ESL model; to meet more tightly the accounting standard further developments of the model should include, at least, a heterogeneous portfolio, the EAD and the LGD that we do not discuss here.

In summary, the ESL model is able to simulate the time series of the Table 1

Sim. Data	for	of
$\mathbf{N}_t$ and $\mathbf{P}_t$	Estimation over $\mathbb{T}_{EST}$	Z_t^{EST}
Y_t and γ_t	Identification over $\mathbb{T}_{IDF}$	Z_t^{IDF}
Y_t and γ_t	Extrapolation over $\mathbb{T}_{EXT}$	Z_t^{EXT}

Table 1: Main series simulated with the ESL model, and series estimated-identified-extrapolated with the procedures of this paper.

where Z_t^{EXT} allows for extrapolation[9] over $\mathbb{T}_{EXT}$ of $\mathbf{N}_t^{EXT}$ and $\mathbf{P}_t^{EXT}$, by means of which the IFRS9-like base-

[9] $\mathbf{N}$ refers to absolute values migration matrices and $\mathbf{P}$ to probabilities.

line is implemented in order to declassify rating classes into a 3-bucket model according to the *segmentation* requirement and to obtain *forward-looking EL* estimates for each bucket, conditionally to a *prospective* macroeconomic scenario: the regulator considers these aspects are baseline foundational principles for a well-balanced, *ex-ante* and forward-looking assessment of bank's credit risk.

Before proceeding further, the main assumptions of the theoretical model are now explicated to specify the estimation problem about the unknown systematic factor Z_t.

As explained with details in [12] and [11, eq.s (12-14)], individual creditworthiness $W_{i,t}$ is assumed to be positively ρ-correlated[10] to an unknown systematic factor Z_t, that it is assumed to explain the *GDP* growth rate γ_t, i.e. $Z_t = Z(\gamma_t) \sim \mathcal{N}_{[-2,+2]}(0,1)$, and to an idiosyncratic term $Z_{i,t}$, i.e. $Z_{i,t} \sim \mathcal{N}(0,1)$. The micro-level model for creditworthiness is therefore (see [11, 12, 8, 13])

$$W_{i,t} = \sqrt{\rho} \cdot Z_t + \sqrt{1-\rho} \cdot Z_{i,t} \Rightarrow Z_{i,t} = \frac{W_{i,t} - \sqrt{\rho} \cdot Z_t}{\sqrt{1-\rho}} \tag{1}$$

According to (1), once the growth rate γ_t is known and $\rho \in [0,+1]$ is fixed (i.e. homogeneously across debtors and constant through time), individual creditworthiness $W_{i,t}$ can be simulated: notice that, the lower ρ the higher the dependence of individual creditworthiness on idiosyncratic shocks. Given a set of bins[11] ℓ_k^h from a known migration matrix, the estimator

$$\widehat{z}_{h,k}(Z_t, \rho) = \frac{\ell_k^h - \sqrt{\rho} \cdot Z_t}{\sqrt{1-\rho}} \tag{2}$$

can be defined for each cell (h,k) of the migration matrix conditionally to the systematic factor Z_t with correlation ρ. Since $p_{h,k} = \Phi(\ell_{k+1}^h) - \Phi(\ell_k^h)$ is the probability of migration into class k from class h then

$$\widehat{p}_{h,k}(Z_t, \rho) = \Phi(\widehat{z}_{h,k+1}(Z_t, \rho)) - \Phi(\widehat{z}_{h,k}(Z_t, \rho)) \tag{3}$$

Therefore, the problem is to estimate ρ and Z_t consistently with the *observed* (i.e. simulated) series of the migration rates matrices $\{\mathbf{P}_t : t \in \mathbb{T}_{EST}\}$. A further problem is to extrapolate the expected evolution of the systematic factor consistently to a prospective scenario for the whole economy. Both issues are addressed to the following section.

3. Micro-simulation modelling procedure

For the ease of readability, Figure 1 reports the main structure of the simulation procedure.

[10]For the ease of a clearer notation, the coefficient β in [11, eq. (14)] here reads as ρ.

[11]See [11, eq.s (19) and (23) and Table 3].

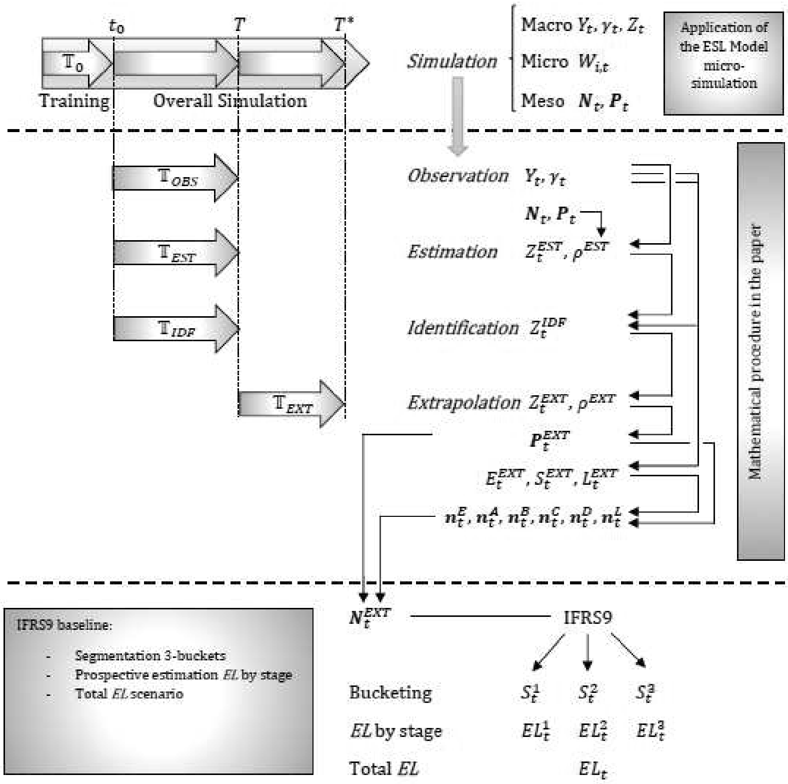


Figure 1: A summary of the procedure. Being our DGP (data generating process), the first section has been simulated as a single run because it generates the micro-level data of the portfolio. The following two sections have been implemented in a Montecarlo algorithm.

In this section we assume the ESL model of [11] has been run[12] so that quantities needed in the following mathematical procedure are usable as if they were observed along $\mathbb{T}_{OBS}$, that coincides with the estimation ($\mathbb{T}_{EST}$) and identification ($\mathbb{T}_{IDF}$) periods ranging from $t_0 + 1$ up to the present T. Dates from $T + 1$ up to T^* define the extrapolation period $\mathbb{T}_{EXT}$ of the future expected dynamics of the quantities involved.

[12] The following results had been obtained by implementing the whole procedure described in the paper in a Montecarlo simulation framework that had been run for 500 iterations. As data are completely simulated, the following results should be considered only for exhibition purposes.

3.1. Estimation

Although inspired to [12], the procedure involved for numerical estimation of ρ and Z_t is different and follows a more numeric-algorithmic approach that can be easily implemented for practical purposes[13].

Let $N_\rho \gg 1$ be an integer and $r \in [0, N_\rho]$ such that the set $\boldsymbol{\rho} = \{\rho_r = \frac{r}{N_\rho} \in [0,1]\}$ defines $R = N_\rho + 1$ candidate values for the correlation ρ that rules creditworthiness (1). Let t_0 be the last date of the training window $\mathbb{T}_0$ and assume having simulated, i.e. as if they were observed, a series $\{\mathbf{P}_t : t \in \mathbb{T}_{EST}\}$ of migration probability matrices that can be obtained by means of the ESL model, remember that $\mathbb{T}_{EST} \equiv \mathbb{T}_{OBS}$. The estimation procedure is as follows:

Step-1 Set $r := r + 1$ and get $\rho_r \in \boldsymbol{\rho}$.

Step-2 Set $t := t + 1 \in \mathbb{T}_{OBS}$ and get $\mathbf{P}_t$.

Step-3 Let $N_Z \gg 1$ be an integer. Generate a random sample $\boldsymbol{\zeta}_{r,t} = \{\zeta^u_{r,t} \sim \mathcal{N}_{[-2,+2]}(0,1) : u \leq N_Z\}$. For the r-th candidate ρ_r and the t-th matrix $\mathbf{P}_t$; a set of N_Z candidate values of Z_t follows, each labelled as $\zeta^u_{r,t}$.

Step-4 Let h be the origin-grade (row) and k the destination-grade (column) identifying the migration cell (h,k) in the matrix. The migration probability $\widehat{p}_{h,k}(\zeta^u_{r,t}, \rho_r) = \Phi(\widehat{z}_{h,k+1}(\zeta^u_{r,t}, \rho_r)) - \Phi(\widehat{z}_{h,k}(\zeta^u_{r,t}, \rho_r))$ follows with (3): repeat $\forall h, k \leq J$ to get the candidate matrix $\widehat{\mathbf{P}}_{r,t,u} = \{\widehat{p}_{h,k}(\zeta^u_{r,t}, \rho_r) : h, k \leq J\}$.

Step-5 $\forall u \leq N_Z$ evaluate $e_{r,t,u} = RMSE(\widehat{\mathbf{P}}_{r,t,u}, \mathbf{P}_t)$ as an overall measure[14] of fitting of the estimated matrix w.r.t. the observed one and get $\mathbf{e}_{r,t} = \{e_{r,t,u} : u \leq N_Z\}$.

Step-6 Find $j : e_{r,t,j} = \arg\min_u \{e_{r,t,u}\}$ and set $Z_{r,t} = \zeta^j_{r,t}$.

Step-7 Repeat from Step-2 until $t = T$ and get $\mathbf{Z}(\rho_r, \mathbb{T}_{EST}) = \{Z_{r,t} : t \in \mathbb{T}_{EST}\}$ as the best performing systematic factor series given ρ_r.

Step-8 Repeat from Step-1 until $r = R$ and get a set of series $\mathbf{Z}(\boldsymbol{\rho}, \mathbb{T}_{EST}) := \{\mathbf{Z}_r : r \leq R\}$.

Step-9 For each of the series evaluate the related variance $v_r = \mathbb{V}[\mathbf{Z}_t]$ and get a vector of variances $\mathbf{v} = \{v_r : r \leq R\}$.

[13]The technique developed in [12] is elegant and mathematically well defined; see also [14]. We did not compare the performance of our procedure to that simply because the systematic factor does not exist in reality and its assumption may be substituted with a different axiom. Nevertheless, as [12] also [11] involves the systematic factor axiom at the root of the modelling and, in absence of other alternatives in the literature, we here maintain it.

[14]$RMSE$ stands for root mean square error.

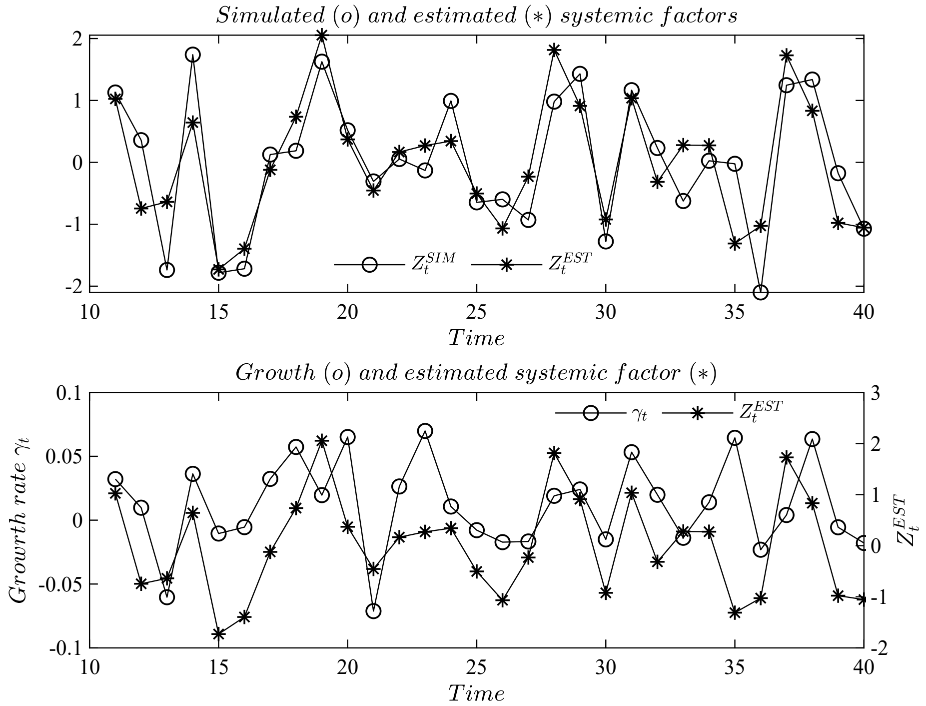


Figure 2: The simulated Z_t^{SIM} and the estimated Z_t^{EST} series of Z_t compared with the growth rate trajectory γ_t along $\mathbb{T}_{OBS} \equiv \mathbb{T}_{EST}$. $Corr[Z_t^{SIM}, Z_t^{EST}] = 0.8166$, $Corr[\gamma_t, Z_t^{EST}] = 0.4338$.

Step-10 By assumption, the systematic factor has a unit variance. Therefore, find $m : v_m = \arg\min_r\{|v_r - 1|\}$ and save $\mathbf{Z}(\rho_m, \mathbb{T}_{EST}) = \mathbf{Z}_m$ as the best-performing series to be candidate of being the estimate of the *unknown* systematic factor series.

The outcome of this procedure is the following: the best-performing estimate of the correlation $\rho^{EST} = \rho_m$ and the best-performing series of the systematic factor $\mathbf{Z}(\rho^{EST}, \mathbb{T}_{EST}) = \{Z_t^{EST} = Z_{m,t} : t \in \mathbb{T}_{EST}\}$ that is consistent with the estimated correlation. Notice that the larger are N_ρ and N_Z the more the estimates are finely tuned. Figure 2 proposes the simulation and estimation outcomes with the following setting[15]: $t_0 = 10$, $T = 40$, $\rho = 0.0163$, $N_\rho = 40$ and $N_Z = 1,000$.

[15]To initialize the micro-level creditworthiness simulation, the ESL model has been calibrated on the migration matrix reported in [12], here reshaped to a master-scale with four grades. The value $\rho = 0.0163$ is the same estimated by [12]. The series $\{Z_t^{EST} : t \in \mathbb{T}_{EST}\}$ is the Montecarlo estimate with 500 iterations of the estimation procedure.

3.2. Identification

This section deals with what we call, with some abuse of terminology, the *identification procedure*, about which few remarks are worth pointing out. Although we use the same term *identification* and we involve simple stochastic equations, it should be safely clear that the described procedure does not concern the well-known *identification problem* in econometrics. To the ends of this paper, *identification* simply means that we aim to estimate the best-performing set of parameters of an equation in order to replicate observed or estimated data just in terms of fitting capability, regardless of any underlying theoretical-economic reasoning. Said differently, behind the identification procedure there are not theoretical assumptions from the economic-financial literature but only numerical fitting capability ambitions, that is more or less what in practitioners often do when asked to solve technical problems, e.g. those the IFRS9 advances. Of course, we are aware that this lack of theoretical assumption may weaken the theoretical foundation of the procedure but, at the same time, we are also as well aware that, at the current stage of available experience, the new regulations make practice more impelling than theory. Therefore, leaving more sound theoretical foundations for a second stage development, we provide a somewhat original and simple *identification* procedure.

Based on the outcomes from the ESL model a series of *observed* (i.e. simulated) series of migration probability matrices is available. This series is consistent with the simulated (i.e. as if it were observable) systematic factor process Z_t, that would be otherwise unknown and not knowable by truth. But, in the previous section, a procedure has been developed to estimate it from a known series of migration probability matrices: in this paper, they have been simulated but a bank may involve a real observed series of migration rates matrices.

One of the innovations introduced by the IFRS9 is that future expected losses should be anticipated by means of a *prospective* scenario. Therefore, as Z_t is entangled to the growth rate γ_t of the GDP Y_t one should link the future evolution of the systematic factor to the dynamics of the GDP. To obtain a *forward-looking* estimate of Z_t conditional to the available macroeconomic scenario about γ_t, a set of equations has been specified to identify the best-performing one that is able to extrapolate the estimates of the systematic factor Z_t^{EXT} in terms of the GDP growth rate: the IFRS9 requires future estimates to be prospective, meaning that they should be conditional to a future macroeconomic scenario[16].

Therefore, let the widest identification period covering the estimation window, i.e. $\mathbb{T}_{IDF} \equiv \mathbb{T}_{EST}$, for

[16]More sophisticated macroeconomic modelling would be worthwhile for further developments, as well as the possibility to shape alternative scenarios to stress the whole procedure.

which the series $\{\gamma_t\}$ is somehow known: in this case, simulated and observed are synonymous. Let also $p > 0$ be an integer indexing different lags and specify the following equation

$$Z_t^{EST} = \alpha_0 + \sum_{j=0}^{p} \beta_j \gamma_{t-j} + \epsilon_t \ : \ \epsilon_t \sim WN(0, \sigma^2) \ , \ t \in \mathbb{T}_{EST} \equiv \mathbb{T}_{IDF} \tag{4}$$

where $\alpha, \beta_j \in \mathbb{R}$. Equation (4) represents a family of stochastic equations, each with its own R_p^2 evaluating the explained variance at the p-th order lag, each with parameters' estimate $\widehat{\boldsymbol{\beta}}_p = (\widehat{\alpha}_0, \widehat{\beta}_0, \widehat{\beta}_1, \ldots, \widehat{\beta}_p)$: as $T = 40$, it has been reasonably assumed that p should not exceed 10, i.e. $0 < p \leq 10$. To find the best-performing equation let $d : R_d^2 = \arg\max_p\{R_p^2\}$ and set $\boldsymbol{\beta}^{IDF} = \widehat{\boldsymbol{\beta}}_d$, then

$$Z_t^{IDF} = \alpha_0^{IDF} + \sum_{j=0}^{d} \beta_j^{IDF} \gamma_{t-j} \tag{5}$$

gives the *identified* model along the identification window $\mathbb{T}_{IDF} = [20, 40] \subset \mathbb{T}_{EST}$[17]. Notice that Z_t^{IDF} is defined as a linear combination of random variables $\{\gamma_t\}$ that, in this paper, have been simulated but, for a more realistic modelling, they can be substituted with real data. A remark is now worth highlighting: as known, the R^2 grows with the number of explanatory variables, hence it is almost evident that the best-performing model should be the one with the highest lag, i.e. $p = 10$. To the end of the purposes of the present paper, we have no ambition to explain the reasons why the systematic factor moves through time, rather we are only interested in numerically replicating at the best its series. On the other hand, the systematic factor is unknown by definition -and not knowable indeed- while, at best, one may have some raw estimation of its realizations without certainty of having correctly made inference about its data-generating-process. Furthermore, it should be safely clear that we are not proposing a theoretically based financial-economic modelling but some practical devices that, in the frame of credit risk migration rates modelling, could be easily implemented to meet some of the practical needs put forth by the IRFS9: to the best of our knowledge this is the first attempt at technically managing the problems arising with the IFRS9 within a simulation framework, hence we decided to keep the modelling as simple as possible.

After this due remark, $\widehat{\boldsymbol{\beta}}_{10} = (\widehat{\alpha}_0, \widehat{\beta}_0, \widehat{\beta}_1, \ldots, \widehat{\beta}_{10})$ is found being the vector of best-performing parameters that allow to specify the identified process $\mathbf{Z}(\mathbb{T}_{IDF} | \boldsymbol{\beta}^{IDF}) = \{Z_t^{IDF} : t \in \mathbb{T}_{IDF}\}$: the outcomes of the identification procedure can be seen in the third panel of Figure 3.

[17] Remember that the first $t_0 = 10$ periods concern the training window, that is neglected, hence the *observed* series start at $t_0 + 1 = 11$ to end at $T = 40$. If a d-lag modelling is considered with $d = 10$ then (5) goes from $t = 20$ onward to $T = 40$.

3.3. Extrapolation

At this stage a brief summary may be worthwhile. A realization of the systematic factor process Z_t as now been estimated as Z_t^{EST} and portrayed in Figure 3.
After estimation, we may consider that Z_t^{EST} is the only available information about the systematic factor driving the dynamics of the migration probability matrices ~~that we know, again, just because we simulated it, but a bank may have its own observed true series of $\mathbf{P}_t$.~~
~~One of the IFRS9 requirements is that expected loss estimates for the future must be *prospective*, actually meaning that the future expectations should be estimated as conditional to some forecasts about the GDP dynamics influencing the future migration probability matrices. As long as we know the Y_t process beyond T, we are in the position to have forecasts about the GDP and its dynamics but, in the real world, a commercial bank may involve forecasts provided by the central bank or any other institution that is accredited for shaping future scenarios. Therefore,~~ To obtain the future dynamics of the systematic factor conditional to the future (i.e. scenario-forecast) GDP growth rate, we have to identify the best-performing equation that is able to replicate the estimated series of the systematic factor in the past: this was the task of the identification procedure developed in the previous Section 3.2. ~~ending with the parameters' estimates $\widehat{\boldsymbol{\beta}}_{10} = (\widehat{\alpha}_0, \widehat{\beta}_0, \widehat{\beta}_1, \ldots, \widehat{\beta}_{10})$, by means of which we identified the process $\mathbf{Z}(\mathbb{T}_{IDF}|\boldsymbol{\beta}^{IDF}) = \{Z_t^{IDF} : t \in \mathbb{T}_{IDF}\}$.~~
Now therefore, assuming the bank trusts the growth rate forecast provided by some accredited institution is equivalent to assume the portion of simulated growth rate series along the extrapolation window, as if it were a provided forecast: hence, $\boldsymbol{\gamma}^{EXT} = \{\gamma_t : t \in \mathbb{T}_{EXT}\}$ is assumed as a given future macroeconomic scenario about the GDP growth rate. Therefore, to *prospectively* extrapolate the future trajectory of the systematic factor conditional to such a forecast $\boldsymbol{\gamma}^{EXT}$, the following (deterministic) model may be involved

$$Z_t^{EXT} = \alpha_0^{IDF} + \sum_{j=0}^{d} \beta_j^{IDF} \gamma_{t-j}^{EXT} \ : \ \forall t \in \mathbb{T}_{EXT} \tag{6}$$

Then, $\mathbf{Z}(\boldsymbol{\gamma}^{EXT}, \mathbb{T}_{EXT}|\boldsymbol{\beta}^{IDF}) = \{Z_t^{EXT} : t \in \mathbb{T}_{EXT}\}$ is the extrapolation of the unknown Z_t along the future window $\mathbb{T}_{EXT}$ beyond T: see the bottom panel of Figure 3, it is conditional to a macroeconomic scenario and consistent to the best-performing *identified* process of the *estimated* systematic factor. Therefore, the IFRS9 prospective scenario requirement can be fulfilled: not only, as it will be discussed lately, but a method also exists to transform the information embedded into the scalar systematic factor realizations to a series of future migration probability matrices that allow the estimation of future expected losses.

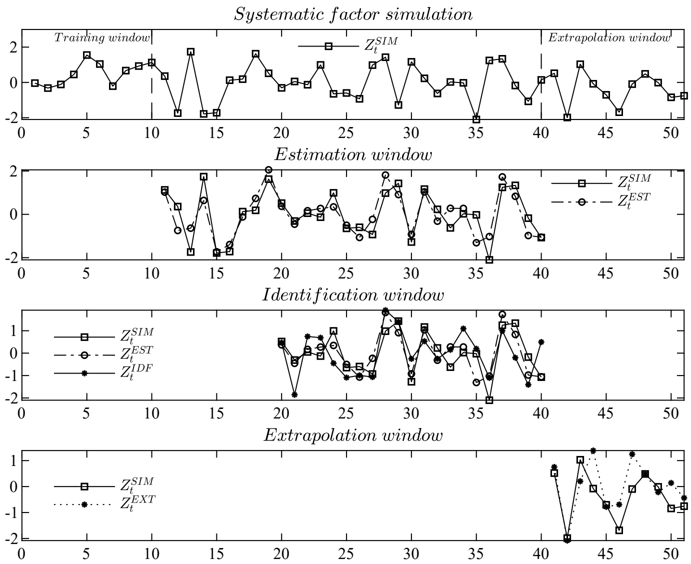


Figure 3: The systematic factor series. First panel: overall simulation, second panel: estimation (see Figure 2)), third panel: identification; fourth panel: extrapolation. Third panel, $t \in \mathbb{T}_{IDF}$: $Corr[Z_t^{SIM}, Z_t^{EST}] = 0.7852$,$Corr[Z_t^{SIM}, Z_t^{IDF}] = 0.5378$, $Corr[Z_t^{EST}, Z_t^{IDF}] = 0.6693$. Fourth panel, $t \in \mathbb{T}_{EXT}$: $Corr[Z_t^{SIM}, Z_t^{EXT}] = 0.7193$.

3.4. Results

This section briefly summarizes what described in previous sections in advance of proceeding with the methodology. Among other quantities, by means of the ESL model developed in [11], a series of simulated data about the systematic factor is available just for control purposes: the top panel of Figure 3 portraits the series of Z_t^{SIM} as if we knew the systematic factor.

The second panel portraits the series of the systematic factor estimated as Z_t^{EST} by means of the procedure developed in Section 3.1: simulated and estimated series are superimposed for a visual direct comparison and, it is once more worth remembering, only Z_t^{EST} is *effectively* known while Z_t^{SIM} it is not known and not knowable beyond simulation.

The third panel of Figure 3 portraits the outcome of the identification procedure developed in Section 3.2: here Z_t^{SIM}, Z_t^{EST} and Z_t^{IDF} are superimposed for direct comparison.

Finally, the bottom panel of the same Figure 3 portraits the extrapolation of a prospective estimate of the systematic factor conditional to a given forecast about the GDP growth rate: this panel superimposes the simulated Z_t^{SIM} data and the extrapolated Z_t^{EXT} data about the unknown systematic factor Z_t.

What can be seen from Figure 3 is that all the portrayed series, within their specific time-windows, are highly and positively correlated. ~~Therefore the mathematical procedure developed in this section turns out being almost secure~~. Clearly, dealing with an *in-silico* laboratory is rather a different situation of having access to real data but, unfortunately, micro-level real data are inaccessible to almost everyone that is not engaged in specific tasks[18] on legal mandate of the regulator or the supervisory authority.

4. Extrapolation of the migration probability matrices

Through the whole simulation window $\mathbb{T}_{SIM}$, a series of migration rates matrices has been simulated by means of the ESL model, namely $\{\mathbf{P}_t : t \in \mathbb{T}_{SIM}\}$ serves for control purposes, as if it were the *true* series available to a bank. Therefore, as $\mathbb{T}_{IDF} \subseteq \mathbb{T}_{OBS} \equiv \mathbb{T}_{EST} \subset \mathbb{T}_{SIM}$, a series of migration rates is usable in each of these periods together with the series Y_t of the GDP and the related growth rate γ_t: along $\mathbb{T}_{OBS} \equiv \mathbb{T}_{EST}$ these are the *known* data and along $\mathbb{T}_{EXT}$ they are *knowable* because some institute has provided forecasts and future macroeconomic scenario.
By means of the estimation and identification procedures discussed above, the systematic factor series has been estimated Z_t^{EST}, identified Z_t^{IDF} and extrapolated Z_t^{EXT}: what is missing is the prospective series of migration rates matrices $\mathbf{P}_t^{EXT}$ conditional to the future macroeconomic scenario.

The correlation coefficient ρ in (1) is assumed constant through time and homogeneous across individuals. As we have estimated the value $\rho^{EST} = 0.0082$, we involve it[19] as suitable also for the future, hence $\rho^{EXT} = \rho^{EST}$. Therefore, having extrapolated Z_t^{EXT} we can now apply (2) as follows

$$\widetilde{z}_{h,k}(Z_t^{EXT}, \rho^{EXT}) = \frac{\ell_k^h - \sqrt{\rho^{EXT}} \cdot Z_t^{EXT}}{\sqrt{1 - \rho^{EXT}}} \tag{7}$$

for all the migrations (h, k) so that (3) now reads as

$$\widetilde{p}_{h,k}(Z_t^{EXT}, \rho^{EXT}) = \Phi(\widetilde{z}_{h,k+1}(Z_t^{EXT}, \rho^{EXT})) - \Phi(\widetilde{z}_{h,k}(Z_t^{EXT}, \rho^{EXT})) \tag{8}$$

[18] None of the here presented series and data is real or somehow based on protected real data.

[19] This estimate sharply differs from the value $\rho = 0.0163$ to which the model has been initialized according to [12]. The reason of this difference is that the here involved migration rates matrix is a recasting of the original 7-grades standard matrix (i.e. with *AAA*, *AA*, *A*, *BBB*, *BB*, *B* and *CCC without D*) into a 6-grades one that includes *A*, *B*, *C and D*, and that as been also *augmented* for entries *E* and exits *L*. Beyond the numerical difference, what is interesting is that involving the *augmented master-scale with default* (i.e. that of a portfolio as an open system with renewal that considers cure events) the correlation increases: this may be due to that the simulated entry-process realistically depends on the GDP growth rate γ_t, to which the systematic factor is entangled to. For a complete development of the ESL simulation model the reader is addressed to [11].

t-1 \ t		E	S: A	S: B	S: C	S: D	L
E		0	p(E,A;t)	p(E,B;t)	p(E,C;t)	p(E,D;t)	p(E,L;t)
S	A	0	p(A,A;t)	p(A,B;t)	p(A,C;t)	p(A,D;t)	p(A,L;t)
	B	0	p(B,A;t)	p(B,B;t)	p(B,C;t)	p(B,C;t)	p(B,L;t)
	C	0	p(C,A;t)	p(C,B;t)	p(C,C;t)	p(C,D;t)	p(C,L;t)
	D	0	p(D,A;t)	p(D,B;t)	p(D,C;t)	p(D,D;t)	p(D,L;t)
L		0	0	0	0	0	p(L,L,t)

Figure 4: The structure of an ESL modelling migrations probabilities with entries (E), exits (L) and grades transitions for stayers (S). Gray-scale or coloured cells refer to stages required by the IFRS9 bucketing developed in Section 6.

and it prospectively extrapolates the migration probability from grade h to grade k at a given $t \in \mathbb{T}_{EXT}$ in the future. By using (8) for all the origin-rows and destination-columns grades, a migration matrix $\mathbf{P}_t^{EXT}$ is obtained as the prospective extrapolation of the migration probability matrix conditional to a future macroeconomic scenario for all $t \in \mathbb{T}_{EXT}$.

For the ease of readability of the following sections, the reader should consider the structure of the migration probability matrix reported in Figure 4 where, in general, $p(h, k; t) = p_{h,k}(Z_t, \rho)$: by setting Z_t^{EXT} and ρ^{EXT} one may extrapolate the migration probabilities $\widetilde{p}_{h,k}$ in (8).

Since the main purpose of the paper is to combine the ESL model for migration rates modelling as open systems with renewal and the simplified IFRS9-baseline, some migration rates deserve particular attention. Therefore, in the following the entry/exit rates and the cure/default rates are discussed.

4.1. Entry and exit rates by grades

By means of (8) we can now prospectively extrapolate the time series $\{\mathbf{P}_t^{EXT} : t \in \mathbb{T}_{EXT}\}$, whose generic element $\widetilde{p}_{h,k}(Z_t^{EXT}, \rho^{EXT})$ provides the future estimate of the migration probability from grade h to grade k, conditionally to the future macroeconomic scenario that is embedded into the forecast by means of the systematic factor extrapolation Z_t^{EXT} obtained with (6), consistently with the correlation ρ^{EXT} obtained with the procedure of Section 3.1. It is worth reminding that this fulfils the IFRS9 baseline concerning the conditioning of forecasts to a prospective macroeconomic scenario.

The augmented master-class Λ_0 described in Section 1 induces a number migration rates. As the ESL model is open to entries and exits we firstly describe them briefly.

Entries refer to row E of the matrix in Figure 4. Once a new position enters the portfolio is not an easy task deciding to what grade it should be addressed. Nevertheless, it comes with a given creditworthiness and the ESL procedure developed in [11, Section 3.1.] describes ~~how the ESL model deals~~ to deal with these events: clearly, a real bank has appropriated and sophisticated procedures to assign the correct rating class to enters. Therefore, each new debtor is placed in class A, B, C or D[20]: specifically, in the real world, a *migration* (E, D) happens in cases of mergers, i.e. when a bank acquires the portfolio of another bank that has estimated the debtor *at-default*. The left panel of Figure 5 reports both the simulated (as if they were observed) and extrapolated trajectories of entry probabilities $\widetilde{p}_{E,k}$ for $k \in \{A, B, C, D\}$: clearly, $\widetilde{p}_{E,E} = \widetilde{P}(E|E)$ is always zero[21] while $\widetilde{p}_{E,D} = \widetilde{P}(D|E)$ defines the probability of a new entry immediately rated *at-default*, that is usually due to mergers. ~~As shown, the extrapolation procedure nicely fits the simulated data along the extrapolation window.~~

Exits refer to column L of the matrix in Figure 4. Basically, a debtor exits the portfolio for the following reasons: she has paid back her commitments regularly, therefore we call $\widetilde{p}_{h,L} = \widetilde{P}(L|h)$ as *regular exits* from *in-bonis* grades $h \in \{A, B, C\}$; she has not paid back her commitments regularly, therefore we call $\widetilde{p}_{D,L} = \widetilde{P}(L|D)$ exit rates from *default*, these rates will almost surely realize losses for the bank; not only, there can be cases of exit right after the entrance, these cases are not reported as they are almost null, $\widetilde{p}_{E,L} = \widetilde{P}(L|E) \approx 0$, indeed these are the cases of those contracts that expire before the *reporting date*; clearly, $\widetilde{p}_{L,L} = \widetilde{P}(L|L) \equiv 1$ as exits accumulate, see [11, Section 3.3] and previous footnote 21. The right panel of Figure 5 emphasizes the exit probabilities with respect to the simulated (i.e. observed) ones. But this is not precisely a drawback, especially in the case of exit from default. Indeed, overestimating these risky exits the procedure simulates the case of a prudential bank.

In the end, although the modelling has been developed within the purely simulation framework of the ESL model, there is evidence for the extrapolation procedure to be realistic enough while, at the same time, being compliant with the IFRS9 baseline about the prospective scenario extrapolation.

[20]Notice that in [11] the class C was absent but it almost straightforward to enlarge the master-scale.

[21]Also $\widetilde{p}_{h,E} = 0$ for all $h \in \{A, B, C, D\}$: a position that is already in the portfolio cannot enter *anew* in the system. In the same way, $\widetilde{p}_{L,E} = 0$ as the entrance of a leaver would be the case of an *old* debtor of the bank that is becoming an *anew* one: from the point of view of the bank this is precisely the same of an entry, no matter if the debtor is an old or a new one, therefore it is managed in the first row, not in the last. As a consequence, and for the sake of completeness, it should be noticed that $\widetilde{p}_{h,L} = 0$ for all $h \in \{A, B, C, D\}$ while $\widetilde{p}_{L,L} = 1$, at each t, because the number of leavers may only accumulate through time in the database of the bank, sharply according to a pure-birth or Poisson process mechanism.

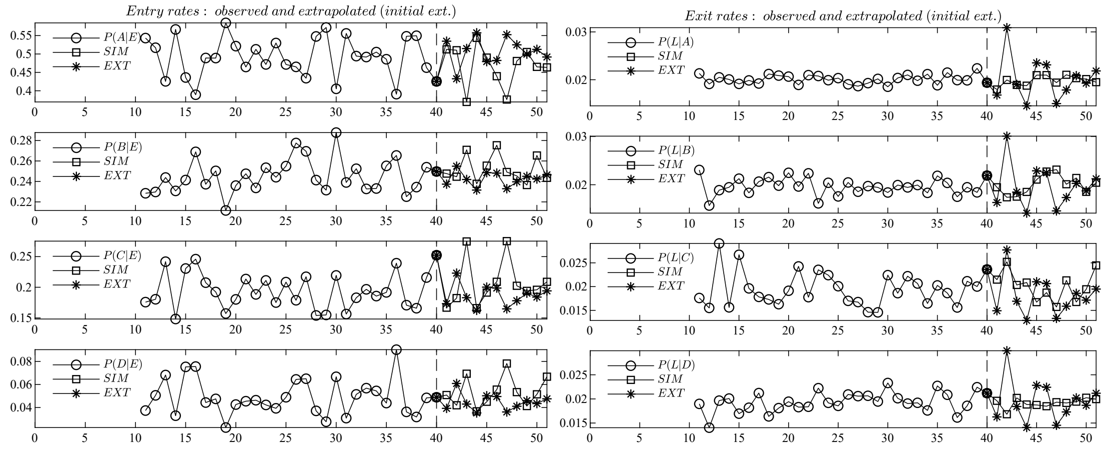


Figure 5: Simulated and extrapolated migration probability time series for new entries (left) and exits (right) by grades on the augmented master-class Λ_0. Entry rates correlation along $\mathbb{T}_{EXT}$ between simulated and extrapolated series, left panel: $(A|E)$: -0.1469, $(B|E)$: 0.2170, $(C|E)$: 0.1941, $(D|E)$: -0.2782. Exit rates correlation along $\mathbb{T}_{EXT}$ between simulated and extrapolated series, right panel: $(L|A)$: 0.4449, $(L|B)$: -0.1846, $(L|c)$: 0.3931, $(L|D)$: -0.6599.

4.2. Cure and default rates by grades

The cure rate $\widetilde{p}_{D,k}$ is the probability for a debtor *at-default* to come back *in-bonis*, that is with $k \in \{A, B, C\}$: see row D from column A to C in the matrix of Figure 4. The standard practice is to set all such probability at zero, i.e. $\widetilde{p}_{D,k} = 0 : k \in \{A, B, C\}$, while setting $\widetilde{p}_{D,D} = 1$ as if it were an absorbing state. As discussed in [11] this is not only inconsistent with reality but it also irremediably weakens the possibility of projection of the migration probability matrix with standard practices in the future as, being D absorbing, in a few steps the whole matrix would collapse in D and the bank would be doomed to fail because almost all of its capital should be put at provision. Clearly, cure events are very a few but nevertheless observed by banks: row D cannot be a vector like $(0, 0, 0, 1)$.

Of course, as it can be easily understood, it is much more likely for a debtor in D to come back in C rather than in B or A. Therefore, $\widetilde{p}_{D,A} \leq \widetilde{p}_{D,B} \leq \widetilde{p}_{D,C}$ should hold for the modelling to be realistic and consistent.

The left panel of Figure 6 portraits the observed (i.e. simulated) and the prospective extrapolation of future cure rates that the bank should expect conditionally to the macroeconomic scenario. Few aspects are worth mentioning: the consistency criterion $\widetilde{p}_{D,A} \leq \widetilde{p}_{D,B} \leq \widetilde{p}_{D,C}$ is fulfilled at each date; the extrapolation procedure underestimates cure rates as if the bank were realistic but more than cautious without being pessimistic, that is a good attitude for a sound *ex-ante* assessment of credit-risk; the extrapolation procedure provides cure rates nicely correlated to the *true* ones that have been simulated with the ESL model for control purposes.

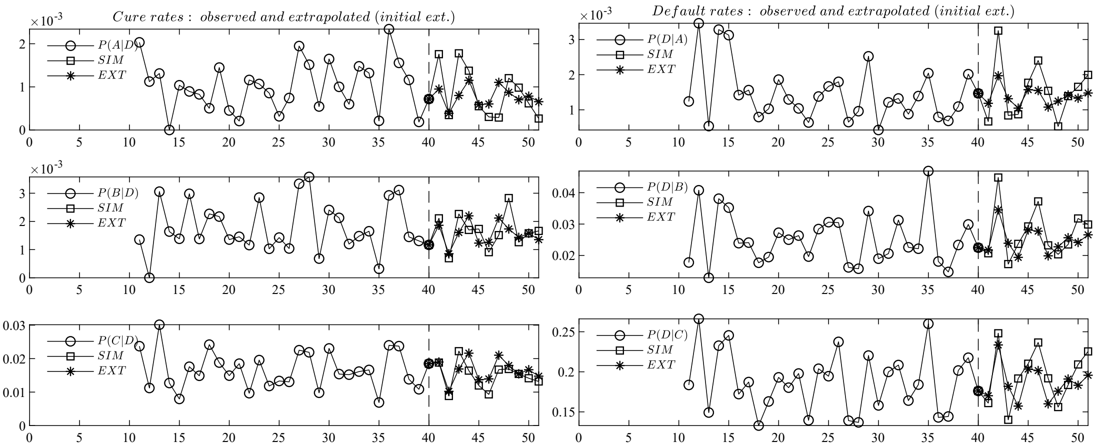


Figure 6: Simulated and extrapolated cure rates time series for new entries (left) and default rates (right) by *in-bonis* grades. Cure rates correlation along $\mathbb{T}_{EXT}$ between simulated and extrapolated series, left panel: $(A|D) : 0.4685$, $(B|D) : 0.5155$, $(C|D) : 0.7047$. Default rates correlation along $\mathbb{T}_{EXT}$ between simulated and extrapolated series, right panel: $(D|A) : 0.8400$, $(D|B) : 0.8122$, $(D|C) : 0.6616$.

With $h \in \{A, B, C\}$ the *default* rate $\widetilde{p}_{h,D}$ is the probability for a debtor *in-bonis* to go *at-default*: see column D from row A to C in the matrix of Figure 4. Said differently, column D reports the probability of default *PD* that, as discussed in sections 1 and 2, is essential to evaluate the *expected loss EL* .

The default rates are ranked with an argument that is complementary to that explained above to rank the cure ones. That is, the higher the quality of the grade of origin the lower the probability to go *at-default*. Therefore, $\widetilde{p}_{A,D} \leq \widetilde{p}_{B,D} \leq \widetilde{p}_{C,D}$ is the criterion of realism and consistency[22].

Let us now point out few remarks concerning the outcomes reported in the right panel of Figure 6. As it can be seen, the criterion of realism and consistency is fulfilled at each date. Since the extrapolation is concerned with a prospective expectation conditional to some forecasts of the macroeconomic scenario, for a bank it is better being a little more prudential than less, of course without being prudential beyond some reasonable threshold. The extrapolated series of the default rates are well correlated with the *true* (i.e. simulated) ones. Therefore, also in the cases of cure and default rates, the extrapolation procedure performs realistically.

[22]Of course, in the real world migrations from very high quality grades to D may happen but the motivations may loosely concern with the creditworthiness *per-se*: said differently, such a dramatic downgrades are reflected by creditworthiness but the causes may come from shocks and reasons beyond the financial-economic performance of the debtor, and beyond the scope of this paper as well.

5. Absolute value migration matrices

As it will be explained with details in Section 6, beyond the *prospective* extrapolation, the IFRS9 also requires the *segmentation* of the portfolio into buckets, i.e. some states should be aggregated with a given rationale, that aims at assessing credit-risk in a realistic and well-balanced manner. A bank we would have time series of data for each exposition, namely a not-balanced panel of debtors[23], and, applying the IFRS9 requirements, one could segment the portfolio into buckets trying to forecast the future evolution of the buckets system, somehow conditionally to a given macroeconomic scenario. As a matter of fact, in this paper we are in the position of an *in-silico* bank, as we have simulated data by means of the ESL model, therefore we could segment our portfolio into buckets and provide the forecasts: from here onward it should be more clear that the ESL model plays the role a laboratory plays to the scientist.

Nevertheless, as long as we simulate a portfolio with a very simplified type of positions, and as long as our aim is to develop an approach that combines the migration matrices modelling as open systems with renewal and some of the baseline principles of the IFRS9, we proceed at the meso-level of analysis, that is we consider groups of positions classified by grades.

One of the original features of the here developed methodology is to embed the open system with renewal approach for migrations rates within the current accounting standard. Previous sections explained how to deal with this issue regarding the prospective extrapolation of migration rates matrices but, to introduce the IFRS9 bucketing principles, some aggregation is needed. Unfortunately, one cannot aggregate states inside the probability matrices compliantly with the IFRS9 because migration probabilities are not addictive compliantly with the bucketing requirement: cells colours in the matrices of figures 4 and 7 give hints about the aggregation rationale that will be discussed in Section 6.

Since now we make clear that operating at the meso-level of the migration rates modelling requires a segmentation that is different to the IFRS9 one, that operates at the micro-level of individual expositions; nevertheless, the developed modelling is logically consistent with the principles of this standard. Therefore, in order to maintain the open system migration rates modelling approach we should firstly transform the extrapolated probability matrices into absolute values matrices, then cells may be aggregated. Shortly, to conjugate the ESL modelling with a IFRS9 baseline, one needs of absolute value migration matrices like the one of Figure 7, where each cell counts how many migration events happen between $t-1$ and t, entries and exits included: namely, each cell $N(h,k;t)$ is a state *occupation number* and all together represent the

[23] The system is indeed open to entries (E) and exits (L), not only stayers (S).

t-1 \ t		E	S: A	S: B	S: C	S: D	L
E		0	N(E,A;t)	N(E,B;t)	N(E,C;t)	N(E,D;t)	N(E,L;t)
S	A	0	N(A,A;t)	N(A,B;t)	N(A,C;t)	N(A,D;t)	N(A,L;t)
	B	0	N(B,A;t)	N(B,B;t)	N(B,C;t)	N(B,C;t)	N(B,L;t)
	C	0	N(C,A;t)	N(C,B;t)	N(C,C;t)	N(C,D;t)	N(C,L;t)
	D	0	N(D,A;t)	N(D,B;t)	N(D,C;t)	N(D,D;t)	N(D,L;t)
L		0	0	0	0	0	N(L,L,t)

Figure 7: The structure of an ESL modelling absolute value migrations with entries (E), exits (L) and grades transitions for stayers (S).

configuration of the system evolving through time.

Therefore, the topic of this section is to prospectively estimate a series of absolute value migration matrices $\{\mathbf{N}_t : t \in \mathbb{T}_{EXT}\}$, i.e. conditionally to a given macroeconomic scenario. The developed procedure involves three phases:

Phase-1 *The whole*. Extrapolation of expectations about entries $\mathbb{E}_t$, stayers $\mathbb{S}_t$ and leavers $\mathbb{L}_t$ groups for the accounting system, the expected total number of contracts the future evolution of the portfolio.

Phase-2 *The parts*. Partitioning the amounts of expected contracts into grades of the master-scale Λ, i.e. $N_{A,t}$, $N_{B,t}$, $N_{C,t}$ and $N_{D,t}$;

Phase-3 *The cells*. Estimating the cells configuration of the matrix in Figure 7, i.e. $N_{h,k,t}$ with $h, k \in \Lambda$.

Notice that this procedure is based on the prospective extrapolation of migration rates matrices as open systems, therefore it conjugates the ESL modelling with simplified IFRS9-baseline principles while overcoming the limitations of the standard procedure discussed in [11].

5.1. Phase Phase-1: the whole

To extrapolate a series $\mathbf{N}_t^{EXT} = \{N_{h,k,t}^{EXT}\}$ of matrices in the future one should firstly extrapolate the volumes of contracts expected to open E_t^{EXT}, persist S_t^{EXT} and close L_t^{EXT} at each $t \in \mathbb{T}_{EXT}$: for a motivation see [11, Eq.s (7) and (8)].

To estimate these terms we follow the same procedure involved for the identification of the systematic factor. As new entries depend (or can be assumed depending) on the cycle of the GDP, consistently with the ESL modelling, here E_t is assumed to be a function of the prospected GDP growth rate of a future macroeconomic scenario, therefore a stochastic equation is specified as follows

$$E_t = \alpha_0^E + \sum_{j=0}^{p_E} \beta_j^E \gamma_{t-j} + \epsilon_t^E \tag{9}$$

where the integer $p_E > 0$ identifies different lags: $\alpha_0^E, \beta_j^E \in \mathbb{R}$.

In advance of any economic-theoretical reasoning, statistical analysis on several simulation outcomes suggested considering that the volume of stayers depends more on the GDP trend than on the related growth rate. Accordingly, a second equation is specified with different integer lags $p_S > 0$

$$S_t = \alpha_0^S + \sum_{j=0}^{p_S} \beta_j^S Y_{t-j} + \epsilon_t^S \tag{10}$$

where $\alpha_0^S, \beta_j^S \in \mathbb{R}$. Finally, according to [11, Eq. (5)], the number of exits is defined as the difference of the *yesterday* volume of contracts in the portfolio and the volume of those who stay in the system for *today*: that is a *flow* quantity. But since the volume of leavers L_t in the ESL modelling accumulates time by time as a *stock* quantity, just like it can be recorded in a bank database, the flow of leavers for the current period is the first order difference $\Delta L_t = L_t - L_{t-1}$. Statistical analysis suggested that such a flow quantity is better explained by the GDP trend rather than the related growth rate. Therefore, the following econometric equation has been specified with different integer lags $p_L > 0$

$$\Delta L_t = \alpha_0^L + \sum_{j=0}^{p_L} \beta_j^L Y_{t-j} + \epsilon_t^L \tag{11}$$

where $\alpha_0^L, \beta_j^L \in \mathbb{R}$. Each econometric equation (9-11) represents a set of equations each involving different lags $p_E, p_S, p_L > 0$, and residual terms are assumed to be white-noise: $\epsilon_t^E, \epsilon_t^S, \epsilon_t^L \sim WN(0, 1)$. Therefore, to choose the best-performing equation in each set we used the R^2 simple criterion: i.e., the best-performing equation in each set is the one whose lag fulfils $d_m = \arg\max_{p_m}\{R^2_{p_m}\}$ for $m \in \{E, S, L\}$ along $\mathbb{T}_{EST}$, that is along the observation window, indeed $\mathbb{T}_{EST} \equiv \mathbb{T}_{OBS}$.

Once parameters $\widehat{\alpha}_0^m$ and $\widehat{\boldsymbol{\beta}}^m$, $m \in \{E, S, L\}$, are identified for the best-performing equations, the available future macroeconomic forecasts of Y_t and γ_t can be involved to extrapolate the expected number of entries

$$E_t^{EXT} = \widehat{\alpha}_0^E + \sum_{j=0}^{d_E} \widehat{\beta}_j^E \gamma_{t-j} \tag{12}$$

and exits

$$L_t^{EXT} = \widehat{\alpha}_0^L + \sum_{j=0}^{d_L} \widehat{\beta}_j^L Y_{t-j} + L_{t-1}^{EXT} \tag{13}$$

With regard to the volume of stayers, a first-candidate extrapolation follows the same rationale

$$\widehat{S}_t = \widehat{\alpha}_0^S + \sum_{j=0}^{d_S} \widehat{\beta}_j^S Y_{t-j} \tag{14}$$

Few remarks are now worth pointing out. First of all, it should be noticed that (12-14) are deterministic as they *define* the extrapolated quantities as linear combinations of random variables provided by some institute that is accredited for shaping future macroeconomic scenarios. Secondly, they operate along the extrapolation window $\mathbb{T}_{EXT}$ for which the future scenario is available. Thirdly, the parameters of each equation are identified along the observation window $\mathbb{T}_{OBS}$, that is feasible for parameters estimation as it coincides with $\mathbb{T}_{EST}$. Applying these parameter estimates along the extrapolation window $\mathbb{T}_{EXT}$ means that we are assuming that the structure of the economy would not substantially change in the future, even if the available macroeconomic scenario anticipates some downturn, recession or, even worst, a crisis. Of course, more sophisticated stochastic equations can be involved to overcome these limitations that are consistent with the Lucas' critique [15]: as we are interested in preparing the experimental principle of a methodology, we decided to maintain the modelling at the simplest level of specification while leaving other economic-theoretical aspects for future developments. Finally, it should be noticed that none of the flow quantities above is put in relation with the extrapolated series of the migration probability matrices $\mathbf{P}_t^{EXT}$ but they are only related to the macroeconomic quantities that explain the prospective extrapolation of the systematic factor Z_t^{EXT} by means of which we obtained $\mathbf{P}_t^{EXT}$.

5.2. Phase Phase-2: the parts

For the sake of notation convenience, let us now consider the standard master-scale of grades specified as before $\Lambda = \{\lambda_1 = A, \lambda_2 = B, \lambda_3 = C, \lambda_3 = D\}$. The marginal occupation numbers of stayers in the absolute matrix of Figure 7 can be specified as follows

$$N_{1,+,t} \equiv A_t \ , \ N_{2,+,t} \equiv B_t \ , \ N_{3,+,t} \equiv C_t \ , \ N_{4,+,t} \equiv D_t \tag{15}$$

where $N_{h,+,t} = \sum_{k=\lambda_1}^{\lambda_4} N(\lambda_h, \lambda_k; t)$ is the h-th row-sum of the inner-matrix concerning the migrations of stayers. Also, let us now define the following series in vector notation

$$\mathbf{y}^h = (N_{h,+,t_0}, \ldots, N_{h,+,T}) \ , \ \mathbf{x} = (Y_{t_0}, \ldots, Y_T) \tag{16}$$

for $h = \lambda_1, \ldots, \lambda_4$ and covering the observation/estimation window $\mathbb{T}_{OBS} \equiv \mathbb{T}_{EST}$. Notice that $\mathbf{y}^h$ is known either because we simulated it or because the bank has the own database. Therefore, by means of a simple

OLS regression we can estimate the parameters $\boldsymbol{\beta}^h = OLS(\mathbf{y}^h, \mathbf{x})$ under standard conditions for each $\lambda_h \in \Lambda$, by assuming the series of residuals is a δ-correlated Gaussian white-noise. Let us indicate parameters estimates with $\widetilde{\beta}_0^h$ and $\widetilde{\beta}_1^h$ and extrapolate future grades occupation numbers as follows

$$\widetilde{N}_{h,+,t} = \widetilde{\beta}_0^h + \widetilde{\beta}_1^h Y_t \ , \ \forall t \in \mathbb{T}_{EXT} = \{T+1, \ldots, T^*\} \tag{17}$$

to estimate the expected volumes of contracts conditionally to the future macroeconomic scenario by grade. Once these extrapolations are available one can evaluate a second-candidate for stayers as follows

$$\widetilde{S}_t = \widetilde{A}_t + \widetilde{B}_t + \widetilde{C}_t + \widetilde{D}_t \ : \ t \in \mathbb{T}_{EXT} \tag{18}$$

By computing weights

$$q_t^A = \widetilde{A}_t / \widetilde{S}_t \ , \ q_t^B = \widetilde{B}_t / \widetilde{S}_t \ , \ q_t^C = \widetilde{C}_t / \widetilde{S}_t \ , \ q_t^D = \widetilde{D}_t / \widetilde{S}_t \ : \ t \in \mathbb{T}_{EXT} \tag{19}$$

the first-candidate extrapolations can be proportionally adjusted as follows

$$N_{h,+,t}^{EXT} = \widetilde{N}_{h,+,t} - q_t^h \cdot (\widehat{S}_t - \widetilde{S}_t) \ \Rightarrow \ \sum_{\lambda_h \in \Lambda} N_{h,+,t}^{EXT} = S_t^{EXT} \ , \ t \in \mathbb{T}_{EXT} \tag{20}$$

so obtaining the final extrapolation of stayers S_t^{EXT}. As it can be proved, the whole S_t^{EXT} is now algebraically consistent with the summation of its parts $N_{1,+,t}^{EXT} = A_t^{EXT}$, $N_{2,+,t}^{EXT} = B_t^{EXT}$, $N_{3,+,t}^{EXT} = C_t^{EXT}$ and $N_{4,+,t}^{EXT} = D_t^{EXT}$ that follow from (17).

5.3. Phase Phase-3: the cells

~~Before proceeding further a brief summary is now worthwhile. By means of~~ (12), (20) and (13) ~~the absolute value of expected entries E_t^{EXT}, stayers S_t^{EXT} and leavers L_t^{EXT}, respectively, are now defined as portrayed in Figure~~ 8~~, together with expected volumes of contracts in grades A_t^{EXT}, B_t^{EXT}, C_t^{EXT} and D_t^{EXT} along the extrapolation window $\mathbb{T}_{EXT}$. Differently said,~~ Totals and marginals of the absolute values matrix of migration events represented in Figure 7 are now extrapolated conditionally to a macroeconomic scenario, what is missing is the configuration of migrations: the extrapolation of these occupation numbers is the topic of this section that provides estimates of the migration cells by means of the migration rates extrapolation $\mathbf{P}_t^{EXT}$ obtained so far with (8).

The following procedure allows to allocate aggregated volumes of observations to the grades of the augmented master-class $\Lambda_0 = \{\lambda_0 = E\} \cup \Lambda \cup \{\lambda_5 = L\}$ where $\Lambda = \{\lambda_1 = A, \lambda_2 = B, \lambda_3 = C, \lambda_4 = D\}$.

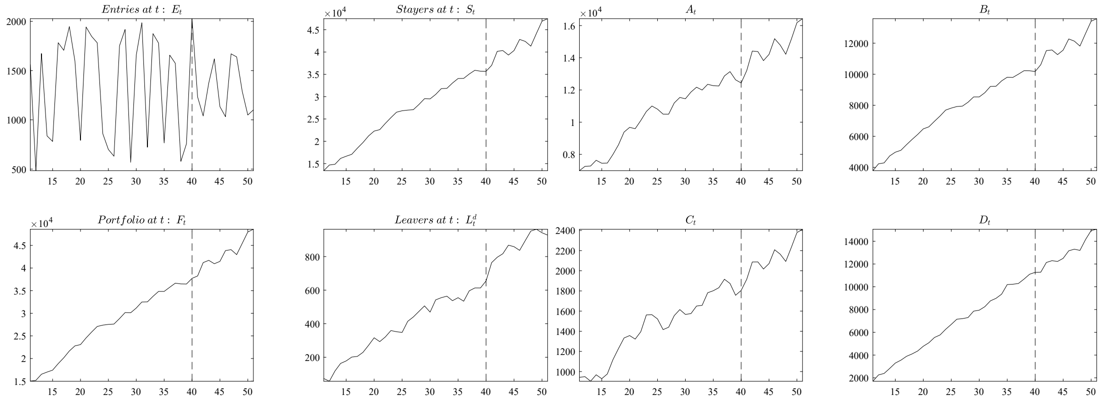


Figure 8: Left four panels: observed and extrapolated series of volumes of entries E_t, stayers S_t, leavers L_t and total volume of contracts F_t in the portfolio. Right four panels: observed and extrapolated series of volumes in grades A_t, B_t, C_t and D_t. The dashed vertical line separates the observed/estimates series along $\mathbb{T}_{OBS} \equiv \mathbb{T}_{EST}$ from the extrapolated ones along $\mathbb{T}_{EXT}$.

Allocation-1 *E: entries into operating grades*. To allocate E_t^{EXT} at each $t \in \mathbb{T}_{EXT}$ into grades consider the first row of the migration probability matrix $\mathbf{P}_t^{EXT}$ excluding the last column concerning leavers while the first column concerning entries is set to zero by construction[24], $\widetilde{p}_{E,E,t}^{EXT} = 0$:

$$\boldsymbol{\pi}_t^E = \left(0, \widetilde{p}_{E,A,t}^{EXT}, \widetilde{p}_{E,B,t}^{EXT}, \widetilde{p}_{E,C,t}^{EXT}, \widetilde{p}_{E,D,t}^{EXT}\right) \Rightarrow \mathbf{n}_t^E = E_t^{EXT} \boldsymbol{\pi}_t^E \tag{21}$$

Allocation-2 *S: stayers migrating over operating grades*. The allocation of S_t^{EXT} into grades follows much the same line of allocating E_t^{EXT} with the difference that, now, we should consider four rows at once. Moreover, as before, we leave aside the last column of leavers while the first one of entries is set to zero by construction[25]

$$\begin{aligned}
\boldsymbol{\pi}_t^A &= \left(0, \widetilde{p}_{A,A,t}^{EXT}, \widetilde{p}_{A,B,t}^{EXT}, \widetilde{p}_{A,C,t}^{EXT}, \widetilde{p}_{A,D,t}^{EXT}\right) \Rightarrow \mathbf{n}_t^A = S_t^{EXT} \boldsymbol{\pi}_t^A \\
\boldsymbol{\pi}_t^B &= \left(0, \widetilde{p}_{B,A,t}^{EXT}, \widetilde{p}_{B,B,t}^{EXT}, \widetilde{p}_{B,C,t}^{EXT}, \widetilde{p}_{B,D,t}^{EXT}\right) \Rightarrow \mathbf{n}_t^B = S_t^{EXT} \boldsymbol{\pi}_t^B \\
\boldsymbol{\pi}_t^C &= \left(0, \widetilde{p}_{C,A,t}^{EXT}, \widetilde{p}_{C,B,t}^{EXT}, \widetilde{p}_{C,C,t}^{EXT}, \widetilde{p}_{C,D,t}^{EXT}\right) \Rightarrow \mathbf{n}_t^C = S_t^{EXT} \boldsymbol{\pi}_t^C \\
\boldsymbol{\pi}_t^D &= \left(0, \widetilde{p}_{D,A,t}^{EXT}, \widetilde{p}_{D,B,t}^{EXT}, \widetilde{p}_{D,C,t}^{EXT}, \widetilde{p}_{D,D,t}^{EXT}\right) \Rightarrow \mathbf{n}_t^D = S_t^{EXT} \boldsymbol{\pi}_t^D
\end{aligned} \tag{22}$$

Allocation-3 *L: leavers from operating grades*. Allocating leavers completes the absolute value migration matrix with the last column that has been previously left aside.

$$\boldsymbol{\pi}_t^{\Delta L} = \left(\widetilde{p}_{E,L,t}^{EXT}, \widetilde{p}_{A,L,t}^{EXT}, \widetilde{p}_{B,L,t}^{EXT}, \widetilde{p}_{C,L,t}^{EXT}, \widetilde{p}_{D,L,t}^{EXT}\right) \Rightarrow \mathbf{n}_t^L = \Delta L_t^{EXT} \boldsymbol{\pi}_t^{\Delta L} \tag{23}$$

[24] No debtor can enter, hence coming from E, into E.

[25] A position that is already in the system cannot enter anew, E.

where $\widetilde{p}^{EXT}_{E,L,t}$ refers to the probability of a new debtor that exits the system before the reporting date and $\Delta L^{EXT}_t = L^{EXT}_t - L^{EXT}_{t-1}$ is the current period exits from the system.

Allocation-4 *Cumulative exits: the last row*. To complete the absolute values matrix, a last row is needed. The elements of this row are all set to zero but the last one that is set to $L^{EXT}_t = \Delta L^{EXT}_t + L^{EXT}_{t-1}$ to account for the cumulative stock of exits from the system up to period $t \in \mathbb{T}_{EXT}$.

The previous four allocations reconstruct the absolute values augmented migration matrix represented in Figure 7 at each date along the extrapolation window in the future. It is now worthwhile highlighting that the so extrapolated matrix fulfils the prospective requirement of the IFRS9 as it depends on the extrapolated series of migration probability matrix $\mathbf{P}^{EXT}_t$, whose elements have been defined in (8). Moreover, it is also consistent with the extrapolated series of expected volumes of entries, stayers and leavers at each date in the future: all these volumes have been estimated as conditional to a given macroeconomic scenario. Therefore, a series $\{\mathbf{N}^{EXT}_t : t \in \mathbb{T}_{EXT}\}$ of extrapolated absolute value matrices is usable and the bucketing is now feasible to meet a segmentation requirement inspired to that of the IFRS9.

6. IFRS9: introducing a baseline

The micro-simulation modelling has now been completed, therefore our *in-silico* bank knows the past history of the portfolio as well as the future extrapolated quantities (i.e. absolute valued migration matrices and their marginals). ~~The simulation window $\mathbb{T}_{STM}$ includes two periods: the observation $\mathbb{T}_{OBS}$, that coincides with the estimation $\mathbb{T}_{EST}$ and identification $\mathbb{T}_{IDF}$ ones, and the extrapolation $\mathbb{T}_{EXT}$ periods. The GDP Y_t and the related growth rate γ_t are said observed for $t \in \mathbb{T}_{OBS}$ while for $t \in \mathbb{T}_{EXT}$ both come from the prospective macroeconomic scenario a given institution may provide: both quantities are therefore known through the whole $\mathbb{T}_{STM}$. Along $\mathbb{T}_{OBS}$ we also know the series of the migration probability matrices, simulated by the ESL model, and along $\mathbb{T}_{EXT}$ we have extrapolated them conditionally to the prospective macroeconomic scenario, moreover we also extrapolated the associated absolute value migration matrices.~~ Therefore, we can now implement the second baseline requirement of the IFRS9: the *segmentation* of the portfolio into *buckets*.

Not depending on the wideness of the augmented master-class Λ, the IFRS9 identifies three buckets for classification (among others see [2], [3], [4], [5], [6]): here, we maintain this structure but, as we are developing a model at the aggregate level, we have to introduce a rather different classification that approximate the rationale of the IFRS9.

Stage-1 This bucket includes positions that did not downgrade between $t-1$ and t: these debtors did not significantly deteriorate their creditworthiness or they improved it significantly while always behaving in *performing* grades. As debtors in this stage are those of best quality, the accounting rationale is to evaluate for them a standard *one-year* $(1Y)$ expected loss (EL). Since the portfolio is a system open to new entries through time, we reasonably assume to Stage-1 should include

(a) all the entries in performing grades (i.e. $N_{E,k,t} : k \in \{A, B, C\}$),

(b) all the positions that maintain for t the same grade they had in $t-1$ (i.e. $N_{h,h,t} : h \in \{A, B, C\}$),

(c) all the positions that improve their creditworthiness from $t-1$ to t in such a way that they upgraded while climbing the master-scale in performing grades (i.e. $N_{B,A,t}, N_{C,A,t}, N_{C,B,t}$). See the matrix structure of Figure 7.

Stage-2 This bucket includes positions that downgraded while being and remaining in a performing grade (i.e. $N_{A,B,t}, N_{A,C,t}, N_{B,C,t}$) or those that cured their *not-performing* grade into a performing one (i.e. $N_{D,k,t} : k \in \{A, B, C\}$): cure events might be more rewarded but, reasonably enough, they are prudentially and generically placed in Stage-2: different banks may have different sensitivity to cures. For all the positions in Stage-2 the accounting requirement is to evaluate a *life-time* (LT) expected loss.

Stage-3 This bucket includes positions gone *at-default*, no matter what the *performing* grade of origin was (i.e. $N_{h,D,t} : h \in \{E, A, B, C\}$) and those that were and still are *not-performing* (i.e. $N_{D,D,t}$). As for those in Stage-2, the accounting requirement for debtors in Stage-3 is to evaluate a *life-time* expected loss.

Summarizing, the volumes of debtors in the buckets are as follows[26]

$$S_t^1 = N_{E,A,t} + N_{E,B,t} + N_{E,C,t} + N_{A,A,t} + N_{B,A,t} + N_{C,A,t} + N_{B,B,t} + N_{C,B,t} + N_{C,C,t} \tag{24}$$

$$S_t^2 = N_{A,B,t} + N_{A,C,t} + N_{B,C,t} + N_{D,A,t} + N_{D,B,t} + N_{D,C,t} \tag{25}$$

$$S_t^3 = N_{E,D,t} + N_{A,D,t} + N_{B,D,t} + N_{C,D,t} + N_{D,D,t} \tag{26}$$

[26]By simplifying assumption the contracts in the simulated portfolio refer to a single financial instrument type, therefore all positions are of the same kind (e.g. mortgages) and we further assume they have the same maturity, for instance, 10 years. This paper aims at introducing an experimental approach, therefore we maintained the modelling at the simplest level. To account for different maturities, say in the number of M, each cell should be decomposed into M^2 sub-cells somehow following the procedure outlined above.

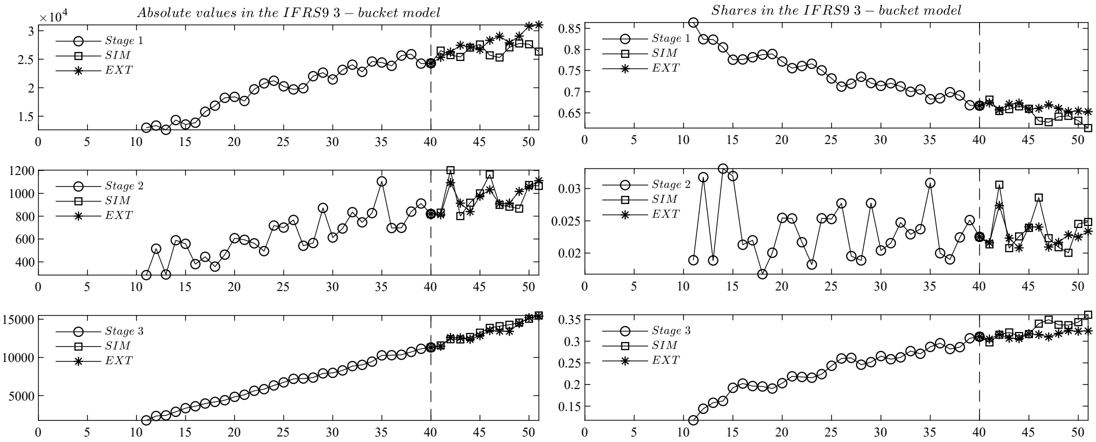


Figure 9: The IFRS9 3stages buckets: absolute values (left) and shares (right). Correlations along $\mathbb{T}_{EXT}$ between simulated and extrapolated series. Absolute values series, left panel: $Stage_1$: 0.1458, $Stage_2$: 0.7718, $Stage_3$: 0.9609. Shares series, right panel: $Stage_1$: 0.6183, $Stage_2$: 0.8070, $Stage_3$: 0.6601.

As it should be now clear, migration probability matrices cannot be aggregated according to the IFRS9 accounting rationale, that collects individual positions according to accounting criteria, but absolute values ones can, and this is the reason why in Section 5 we developed a procedure to extrapolate the absolute values conditionally to a macroeconomic scenario. Said differently, the bucketing into the stages at aggregate or meso-level mimics that of the IFRS9, fulfils either the *prospective* and the *segmentation* criteria, and it is consistent with the *open system migration rates modelling* of the ESL model. Of course, what proposed is a simplified baseline of the IFRS9 principles that include a number of more sophisticated requirements that we do not consider in this paper.

Figure 9 portraits the 3−buckets portfolio into stages along the whole simulation window after the training one (i.e. $\forall t > t_0$) while separating the observation (i.e. $t \in [t_0 + 1, T]$) from the extrapolation window (i.e. $t \in [T + 1, T^*]$): as it can be seen, all the series along $\mathbb{T}_{EXT}$ nicely correlate with those simulated for control purposes. The one described is the case of a portfolio that increasingly deteriorates as the share of positions in Stage-3 increases sharply doubling the concentration of risky positions, while those in Stage-1 decreases, although the former together with positions in Stage-2 amount to about the 70%[27].

Before proceeding further let us describe the configuration of the portfolio in terms of components. Let therefore $x_{k,t} = \sum_{h \in \Lambda} N_{h,k,t}$ be the volume of debtors in grades $k = \{A, B, C, D\}$ at date t, i.e. the vertical

[27] Notice that the presented outcomes are just for exhibition purposes.

summation of the k-th column of the absolute value migrations. In general, the portfolio's configuration is

$$\mathbf{x}_t = (x_{A,t}, x_{B,t}, x_{C,t}, x_{D,t}) \tag{27}$$

where, by definition, $x_{D,t} \equiv S_t^3$ is the only component that can be assigned to a stage unequivocally. Nevertheless, having $N_{h,k,t}^{EXT}$ available $\forall t \in \mathbb{T}_{EXT}$ we can outline a portfolio segmentation according to the bucketing into stages by computing the following components referring to (24), (25) and (26)

$$\begin{aligned}
&Stage-1 \qquad x_{A,t}^1 = \sum_{h \in \{E,A,B,C\}} N_{h,A,t} \;\; , \;\; x_{B,t}^1 = \sum_{h \in \{E,B,C\}} N_{h,B,t} \;\; , \;\; x_{C,t}^1 = \sum_{h \in \{E,C\}} N_{h,C,t} \;\; , \\
&Stage-2 \qquad x_{A,t}^2 = N_{D,A,t} \;\; , \;\; x_{B,t}^2 = \sum_{h \in \{A,D\}} N_{h,B,t} \;\; , \;\; x_{C,t}^2 = \sum_{h \in \{A,B,D\}} N_{h,C,t} \;\; , \\
&Stage-3 \qquad x_{D,t}^3 = \sum_{h \in \{E,A,B,C,D\}} N_{h,D,t},
\end{aligned} \tag{28}$$

so defining the following segmentation

$$\begin{aligned}
&Stage-1 \qquad \mathbf{x}_t^1 = (x_{A,t}^1, x_{B,t}^1, x_{C,t}^1, 0), \\
&Stage-2 \qquad \mathbf{x}_t^2 = (x_{A,t}^2, x_{B,t}^2, x_{C,t}^2, 0), \\
&Stage-3 \qquad \mathbf{x}_t^3 = (0, 0, 0, x_{D,t}^3).
\end{aligned} \tag{29}$$

In the IFRS9 accounting rationale, losses are related to debtors *at-default*. Let now $\delta_D\%$ be the average loss of the exposed capital[28], that is the stock of capital due to debtors *at-default*. According to the prudential rationale of the CRD-CRR (Capital Requirements Directive [16], Capital Requirements Regulation [17]) losses should be estimated on expectation over all grades, both performing and non-performing, not only D. Moreover, loss estimates should be conditional to a macroeconomic scenario. Let us then introduce the prospective-PD, that is the extrapolation of the default probability conditional to a future scenario about the economy, i.e. this is the D column of migration probability matrix (see figures 4 and 7)

$$\boldsymbol{p}_t^D = (\tilde{p}_{A,D,t}^{EXT}, \tilde{p}_{B,D,t}^{EXT}, \tilde{p}_{C,D,t}^{EXT}, \tilde{p}_{D,D,t}^{EXT})' \;\; : \;\; t \in \mathbb{T}_{EXT} = [T+1, T^*] \tag{30}$$

The segmented-prospective and total EL can be finally estimated as follows

$$\begin{aligned}
&Stage-1 \qquad EL_t^{1,EXT} = \delta_D \, \mathbf{x}_t^{1,EXT} \cdot \boldsymbol{p}_t^D, \\
&Stage-2 \qquad EL_t^{2,EXT} = \delta_D \, \mathbf{x}_t^{2,EXT} \cdot \boldsymbol{p}_t^D,
\end{aligned}$$

[28] For the sake of simulation purposes it has been set $\delta_D = 75\%$.

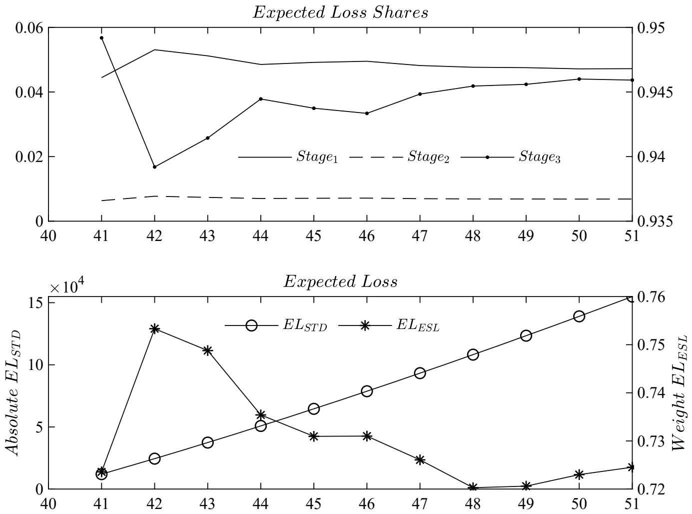


Figure 10: The shares of expected losses in the buckets and the total expected loss compared to standard estimates.

$$
\begin{aligned}
&\textit{Stage}-3 \qquad EL_t^{3,EXT} = \delta_D\, \mathbf{x}_t^{3,EXT} \cdot \boldsymbol{p}_t^D, \\
&\textit{Total} \qquad EL_t^{EXT} = EL_t^{1,EXT} + EL_t^{2,EXT} + EL_t^{3,EXT}.
\end{aligned}
\tag{31}
$$

The top panel of Figure 10 reports the shares of expected loss in each of the buckets according to their stage: the left axis refers to Stage-1 and Stage-2 while the right axis refers to Stage-3. As it can be seen, almost all of the EL is due to positions in Stage-3: being these positions *at-default* this is what we should expect. Therefore, very a little is due to positions in Stage-2 and almost nothing can be attributed to those in Stage-1: the former are positions that performed downgrades, but always remaining in performing grades, or debtors who cured their creditworthiness from grade D to some performing grade; the latter are new or positions that did not perform downgrades or, even more significantly, those who upgraded their creditworthiness. Due to the not so heavy-volatility in the macroeconomic scenario (see the bottom panel of Figure 2) the future expected losses of the buckets are expected evolving almost stationary in the future.

The bottom panel of Figure 10 superposes the series of expected losses estimated according to an almost standard practice (see [11]) and the weight of those estimated with the here developed procedure compared to the standard ones. The standard practice is to estimate only the $1Y-EL$ as if all the debtors were currently in Stage-1 and, to shape some future forecast, a given migration probability matrix $\mathbf{P}_T$ available for the present is projected in the future for some periods after eigen-decomposition

$$
\mathbf{P}_T = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1} \Rightarrow EL_t^{STD} = \mathbf{x}_t \mathbf{V}\mathbf{\Lambda}^t\mathbf{V}^{-1}\ ,\ \forall t \in \mathbb{T}_{EXT}. \tag{32}
$$

As proved by [11] this procedure is doomed to fail if grade D were absorbing and, if it were not, it would inefficiently and dangerously over-estimate the expected loss. The ratio EL_t^{EXT}/EL_t^{STD} in the bottom panel of Figure 10 shows that *a prospective-segmented EL estimation on the basis of an open-sample migration rates modelling is lower than the standard one*: differently said, such estimate is less costly for the bank while being, at the same time, more effective in accounting for an *ex-ante* well-balanced credit risk assessment that accounts for a future macroeconomic scenario about the whole economy. To the ends of this paper this result shows that combining the ESL model with a simplified IFS9-baseline not only goes in the directions traced by the regulator but, also, can be found more efficient in estimating expected losses, so providing a less costly capital provision. Instead of being a definitive result, this finding is better understood as the starting point for further research.

7. Concluding remarks

The main motivation of this paper is to develop a simulation modelling procedure to implement a reduced set of the baseline principles recently introduced with the International Financial Reporting Standard 9 with the aim to estimate Expected Losses in a *prospective* and *segmented* way.

To this end, the ESL model developed in [11] has been here applied as the data generating process for the stochastic evolution of a benchmark portfolio that, under the influence of the macroeconomic dynamics, realistically renews through time with entries and exits while allowing for small number of migrations from default to preforming grades (i.e. cure events).
A Montecarlo procedure has been developed to estimate, identify and extrapolate future dynamics of migration rates matrices consistently with the open-sample structure and conditionally to a future macroeconomic scenario. The simulation outcomes have been compared to those one may obtain with the standard closed-sample approach with the static eigen-decomposition method with absorbing default-state. Compared with this practice, the main result in the paper is that the here developed method turns out being more effective either because it is somehow compliant with some of the current accounting and prudential normative principles, and also because it allows for *a more reliable provision and an ex-ante and forward-looking estimation of expected losses*.

While waiting for the current normative to explicate its effects and for first real data to be available, the modelling has been maintained at the simplest level, hence some aspects have been left for further developments (e.g., heterogeneous portfolios of financial instruments, expected loss as a function of the exposure at default and the loss given default, unexpected losses). This paper may then be considered as a first experimental attempt to develop a methodology to introduce credit risk accounting in economic modelling while considering both the open-sample migration rates approach and the baseline principles of the International Financial Reporting Standard 9 and the prudential capital requirements directive and regulation.

Acknowledgments

The authors are grateful to anonymous reviewers for accurate reports and comments: the quality of the paper has been remarkably improved by following their suggestions. The authors also thank the participants to the MDEF workshop held in Urbino (September 2018) for fruitful discussions, in particular, Laura Gardini, Roberto Dieci, John Barkley Rosser (Jr.) and Anna Maria Grazia Variato. The authors' opinions are their own and do not involve the responsibility of their Institutions. This research was partially supported by IRES Piemonte (Socioeconomic Research Insitute of Piedmont) and by MIUR (Ministero dell'Istruzione, dell'Universitá e della Ricerca Scientifica), Italy. The authors declare no conflict of interest.

References

[1] G20, Declaration on Strengthening the Financial System. London summit, April 2^{nd} 2009 (2009).
URL http://www.g20.utoronto.ca/2009/2009ifi.pdf

[2] BCBE, Guidelines: Guidance on Accounting for Expected Credit Losses., Basel Committee on Banking Supervision; Issued for comment by 30 April 2015, February 2015 (2015).
URL www.bis.org/bcbs/publ/d311.pdf

[3] ECB, SSM thematic review on IFRS 9: Assessment of Institutions' Preparedness for the Implementation of IFRS 9., European Central Bank (2017).
URL www.bankingsupervision.europa.eu/banking/letterstobanks/shared/pdf/2017/ssm.reportlsi_2017.en.pdf

[4] IFRS, IFRS 9 Financial Instruments, Project summary, 24 July 2014., IFRS FOUNDATION (2014).
URL www.ifrs.org/-/media/project/financial-instruments/project-summaries/ifrs-9-project-summary-july-2014.pdf

[5] KPMG, First Impressions: IFRS 9 Financial Instruments. September 2014, KPMG (2014).
URL home.kpmg/content/dam/kpmg/pdf/2014/09/first-impressions-IFRS9.pdf

[6] DELOITTE, IFRS 9 Financial Instruments, IFRS Global Office (2017).
URL www.iasplus.com/en/standards/ifrs/ifrs9

[7] S. Trueck, S. Rachev, Rating Based Modeling of Credit Risk. Theory and Applications of Migration Matrices., 2009. doi:10.1016/B978-0-12-373683-3.X0001-2.

[8] C. Blhum, L. Overbeck, C. Wagner, Introduction to Credit Risk Modeling, Chapman-Hall, 2010.

[9] M. Crouhy, D. Galai, R. Mark, A comparative analysis of current credit risk models, Journal of Banking and Finance 24 (1-2) (2000) 59–117. doi:10.1016/S0378-4266(99)00053-9.

[10] F. Fei, A.-M. Fuertes, E. Kalotychou, Credit rating migration risk and business cycles, Journal of Business Finance and Accounting 39 (1-2) (2012) 229–263. doi:10.1111/j.1468-5957.2011.02272.x.

[11] S. Landini, M. Uberti, S. Casellina, Credit risk migration rates modeling as open systems: A micro-simulation approach, Communications in Nonlinear Science and Numerical Simulation 58 (2018) 147 – 166, special Issue on "Dynamic Models in Economics and Finance". doi:10.1016/j.cnsns.2017.05.028.

[12] L. Forest, B. Belkin, S. Suchower, A one-parameter representation of credit risk and transition matrices, Tech. rep., CreditMetrics Monitor (1998).

[13] M. Gordy, A risk-factor model foundation for ratings-based bank capital rules, Journal of Financial Intermediation 12 (3) (2003) 199–232. doi::10.1016/S1042-9573(03)00040-8.

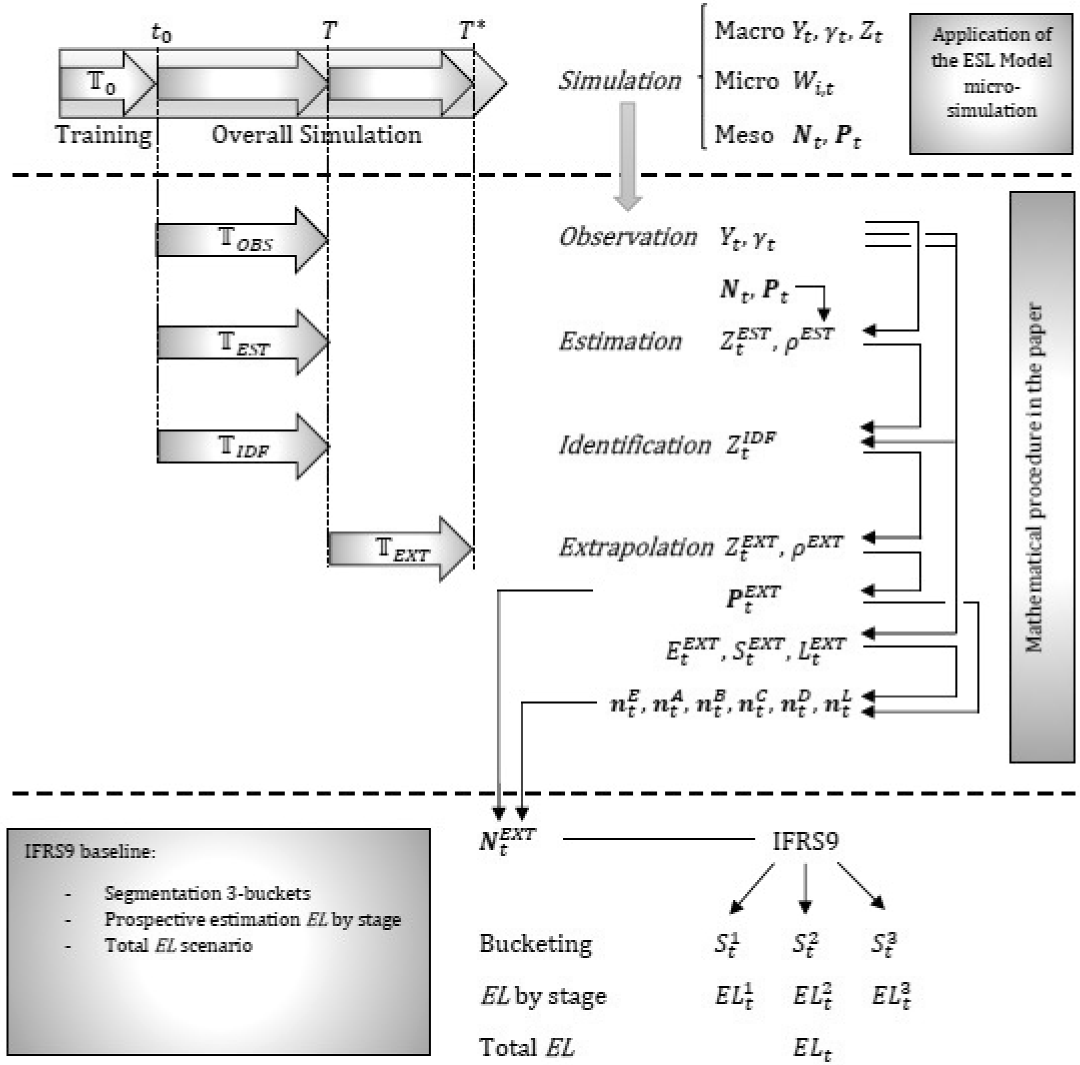
[14] J. Shao, S. Li, Y. Li, Estimation and prediction of credit risk based on rating transition systems, Tech. rep. (2018).
URL arxiv.org/pdf/1607.00448.pdf

[15] R. Lucas, Econometric policy evaluation: A critique, Carnegie-Rochester Conference Series on Public Policy 1 (1) (1976) 19–46.
URL EconPapers.repec.org/RePEc:eee:crcspp:v:1:y:1976:i::p:19-46

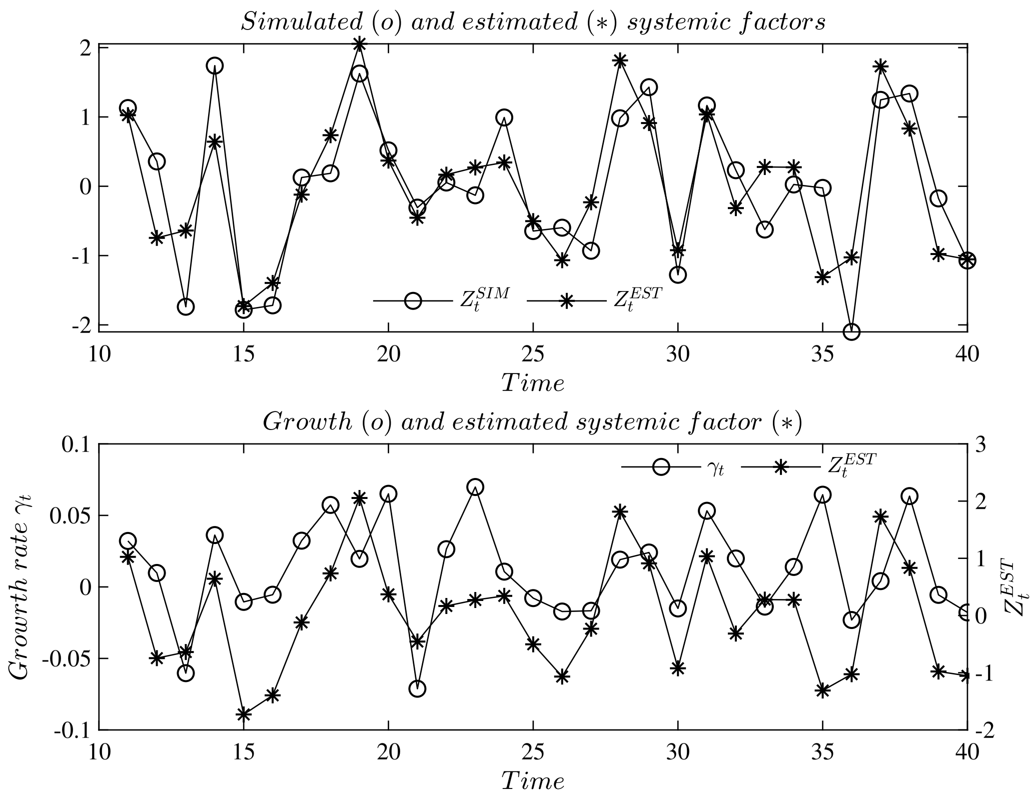
[16] EU, Regulation (EU) No 575/2013 of the European Parliament and of the Council on Prudential Requirements for Credit Institutions and Investment Firms and Amending Regulation (EU) No 648/2012, Official Journal of the European Union (2013).

URL https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32013R0575&from=EN

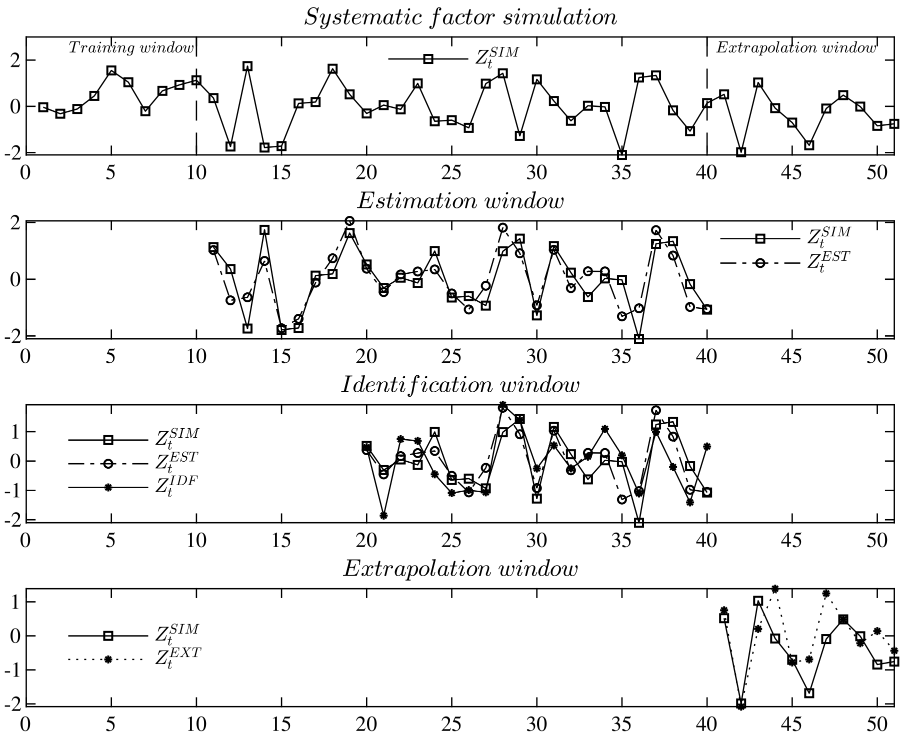
[17] EU, Directive 2013/36/EU of the European Parliament and of the Council on Access to the Activity of Credit Institutions and the Prudential Supervision of Credit Institutions and Investment Firms, Amending Directive 2002/87/EC and Repealing Directives 2006/48/EC and 2006/49/EC, Official Journal of the European Union (2013).
URL https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013L0036&from=EN



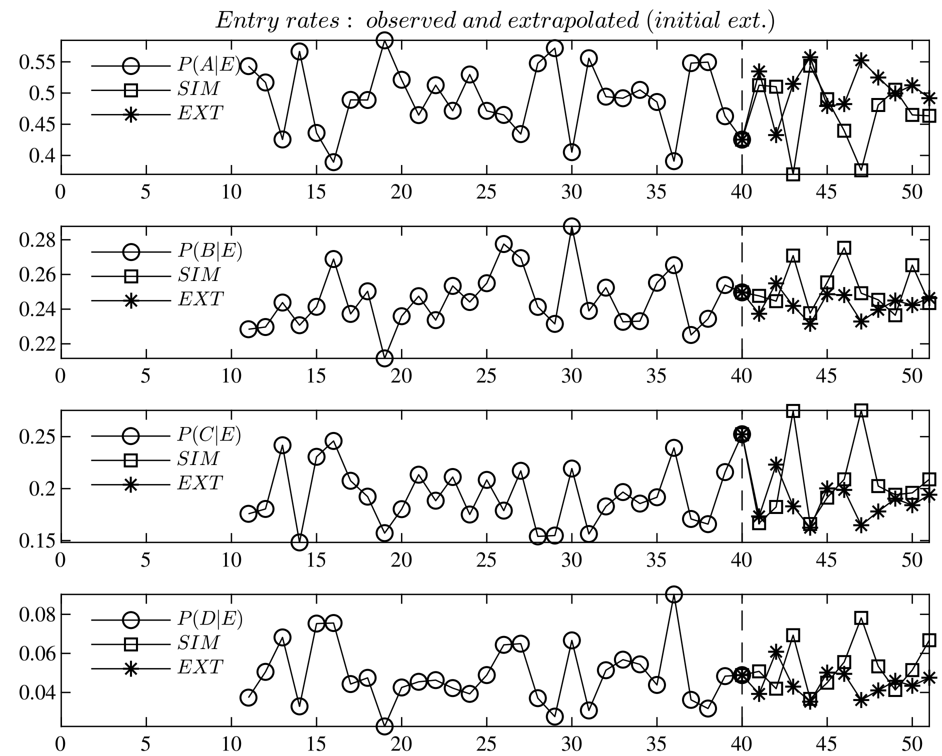
t_0
T
T^*
$\mathbb{T}_0$
Training
Overall Simulation
Simulation
Macro Y_t, γ_t, Z_t
Micro $W_{i,t}$
Meso $\boldsymbol{N}_t, \boldsymbol{P}_t$
Application of the ESL Model micro-simulation
$\mathbb{T}_{OBS}$
$\mathbb{T}_{EST}$
$\mathbb{T}_{IDF}$
$\mathbb{T}_{EXT}$
Observation Y_t, γ_t
$\boldsymbol{N}_t, \boldsymbol{P}_t$
Estimation Z_t^{EST}, ρ^{EST}
Identification Z_t^{IDF}
Extrapolation Z_t^{EXT}, ρ^{EXT}
$\boldsymbol{P}_t^{EXT}$
$E_t^{EXT}, S_t^{EXT}, L_t^{EXT}$
$\boldsymbol{n}_t^E, \boldsymbol{n}_t^A, \boldsymbol{n}_t^B, \boldsymbol{n}_t^C, \boldsymbol{n}_t^D, \boldsymbol{n}_t^L$
Mathematical procedure in the paper
IFRS9 baseline:
- Segmentation 3-buckets
- Prospective estimation EL by stage
- Total EL scenario
$\boldsymbol{N}_t^{EXT}$
IFRS9
Bucketing S_t^1 S_t^2 S_t^3
EL by stage EL_t^1 EL_t^2 EL_t^3
Total EL EL_t



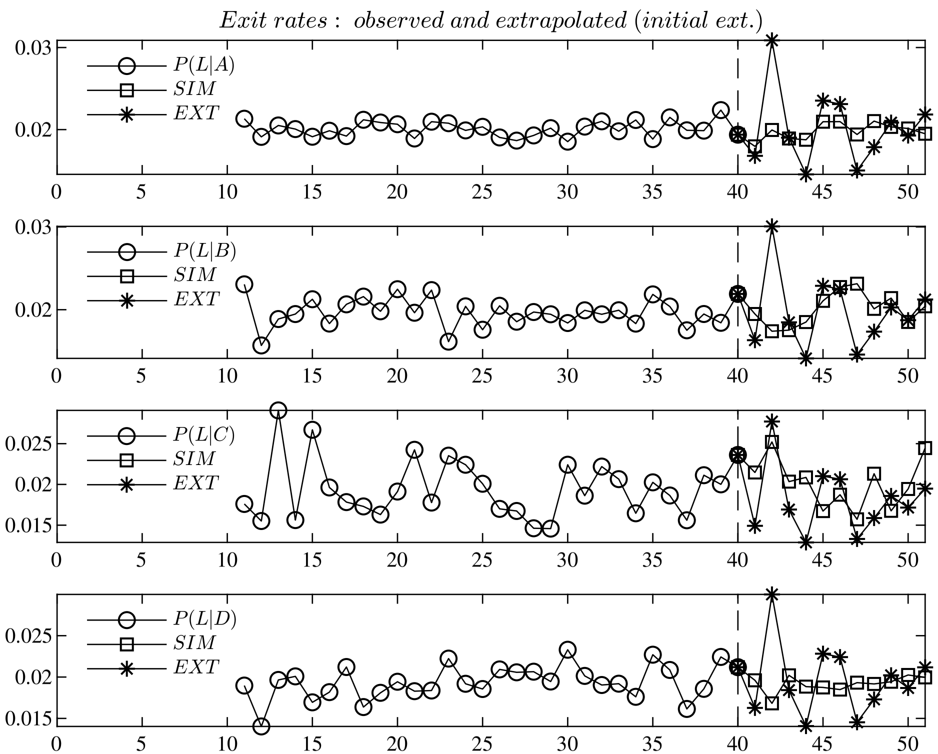
Simulated (o) and estimated (*) systemic factors
Z_t^{SIM}
Z_t^{EST}
Time
Growth (o) and estimated systemic factor (*)
Growrth rate γ_t
γ_t
Z_t^{EST}
Time


Systematic factor simulation
Training window
Extrapolation window
Z_t^{SIM}
Estimation window
Z_t^{SIM}
Z_t^{EST}
Identification window
Z_t^{SIM}
Z_t^{EST}
Z_t^{IDF}
Extrapolation window
Z_t^{SIM}
Z_t^{EXT}

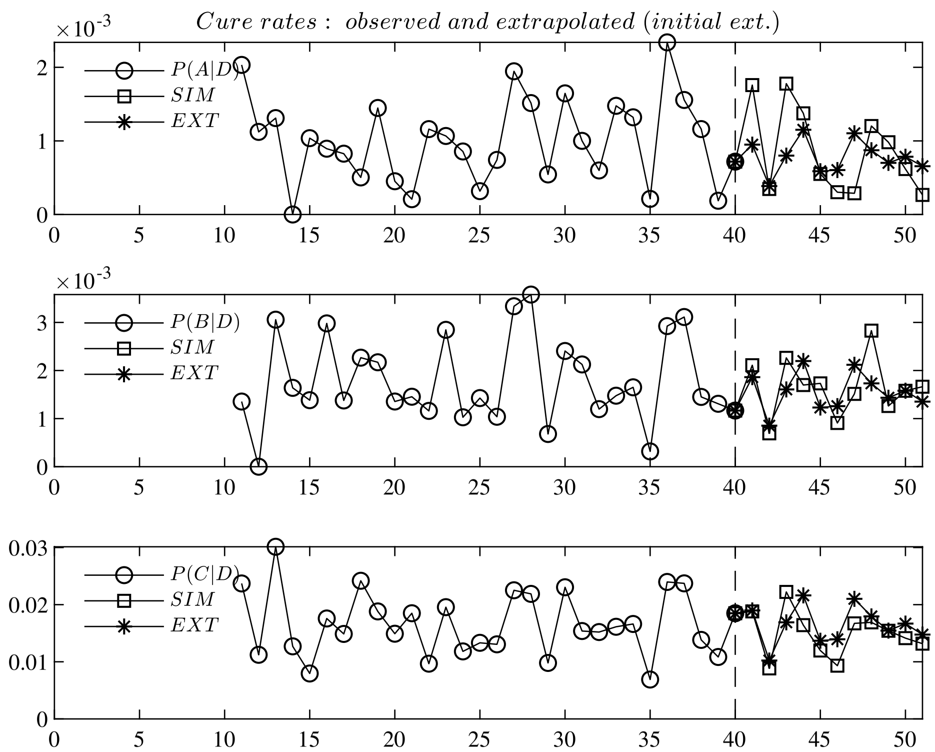
t-1 \ t		**E**	**S**				**L**
			A	B	C	D	
E		0	p(E,A;t)	p(E,B;t)	p(E,C;t)	p(E,D;t)	p(E,L;t)
S	A	0	p(A,A;t)	p(A,B;t)	p(A,C;t)	p(A,D;t)	p(A,L;t)
	B	0	p(B,A;t)	p(B,B;t)	p(B,C;t)	p(B,C;t)	p(B,L;t)
	C	0	p(C,A;t)	p(C,B;t)	p(C,C;t)	p(C,D;t)	p(C,L;t)
	D	0	p(D,A;t)	p(D,B;t)	p(D,C;t)	p(D,D;t)	p(D,L;t)
L		0	0	0	0	0	p(L,L;t)



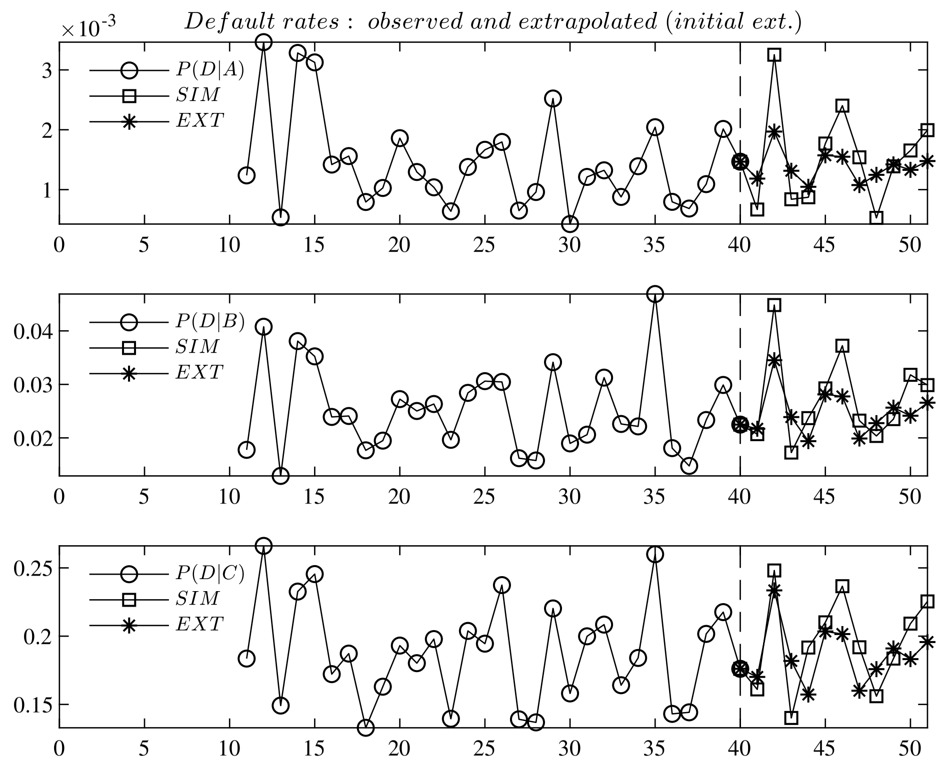
$Entry\ rates:\ observed\ and\ extrapolated\ (initial\ ext.)$
$P(A|E)$
SIM
EXT
$P(B|E)$
SIM
EXT
$P(C|E)$
SIM
EXT
$P(D|E)$
SIM
EXT



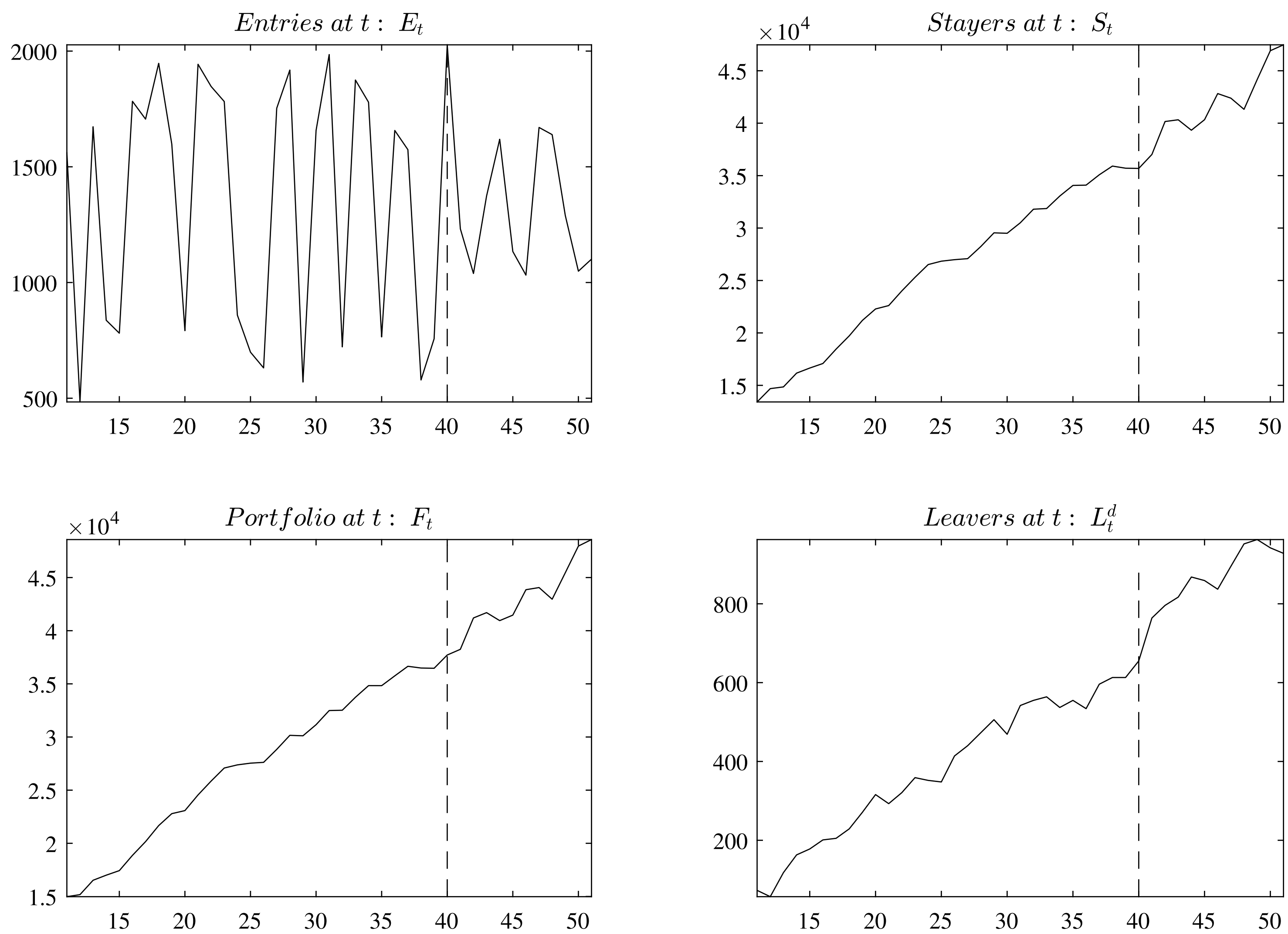
Exit rates : observed and extrapolated (initial ext.)
P(L|A)
SIM
EXT
P(L|B)
SIM
EXT
P(L|C)
SIM
EXT
P(L|D)
SIM
EXT

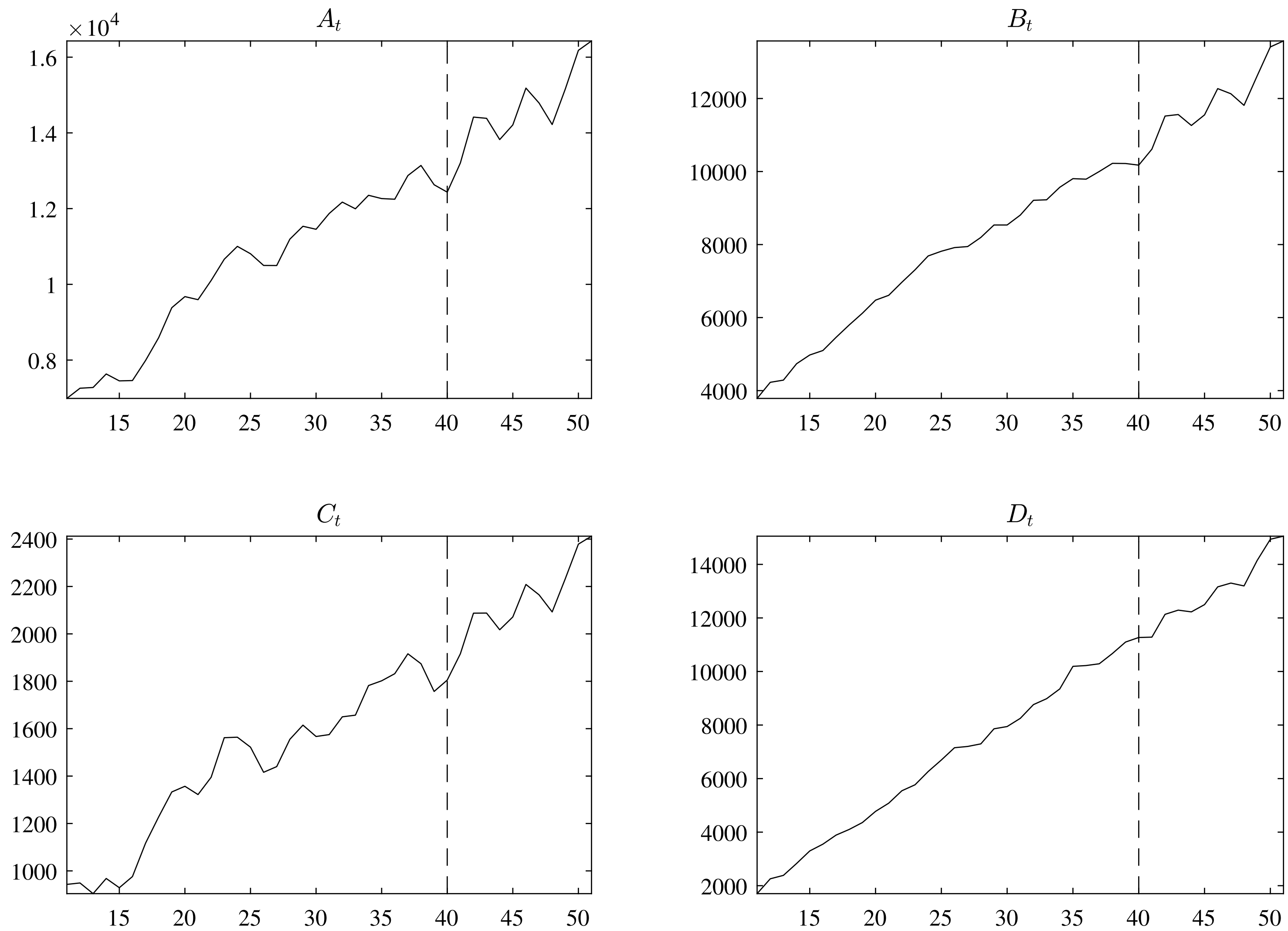


Cure rates : observed and extrapolated (initial ext.)
×10^-3
P(A|D)
SIM
EXT
×10^-3
P(B|D)
SIM
EXT
P(C|D)
SIM
EXT

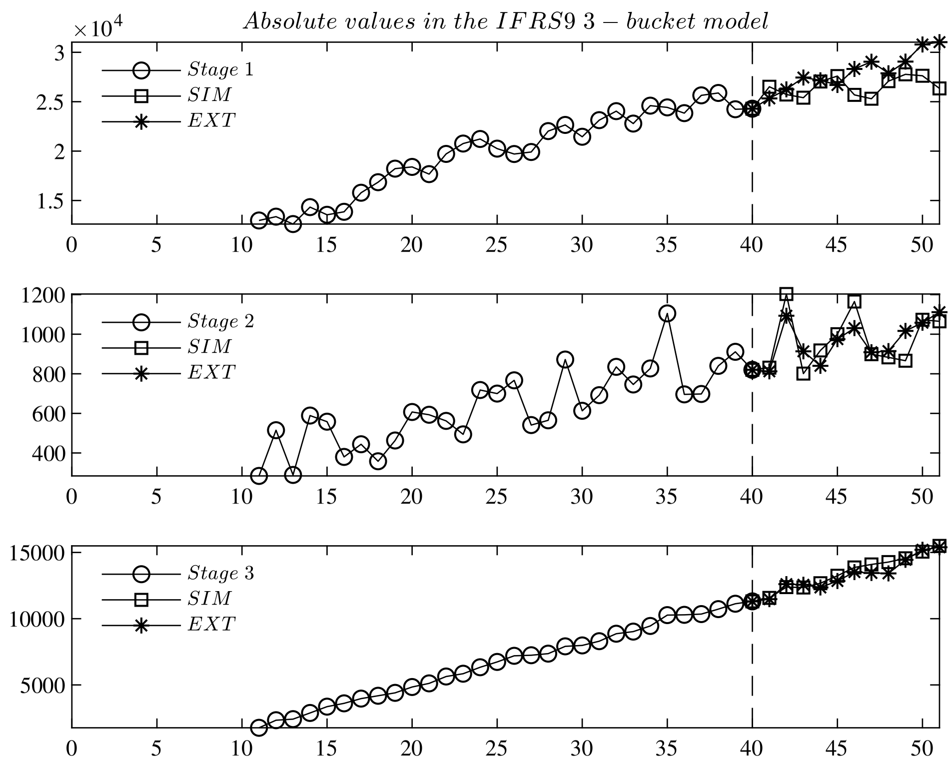

Default rates : observed and extrapolated (initial ext.)
×10^-3
P(D|A)
SIM
EXT
P(D|B)
SIM
EXT
P(D|C)
SIM
EXT

t-1 \ t		E	S: A	S: B	S: C	S: D	L
E		0	N(E,A;t)	N(E,B;t)	N(E,C;t)	N(E,D;t)	N(E,L;t)
S	A	0	N(A,A;t)	N(A,B;t)	N(A,C;t)	N(A,D;t)	N(A,L;t)
	B	0	N(B,A;t)	N(B,B;t)	N(B,C;t)	N(B,C;t)	N(B,L;t)
	C	0	N(C,A;t)	N(C,B;t)	N(C,C;t)	N(C,D;t)	N(C,L;t)
	D	0	N(D,A;t)	N(D,B;t)	N(D,C;t)	N(D,D;t)	N(D,L;t)
L		0	0	0	0	0	N(L,L,t)

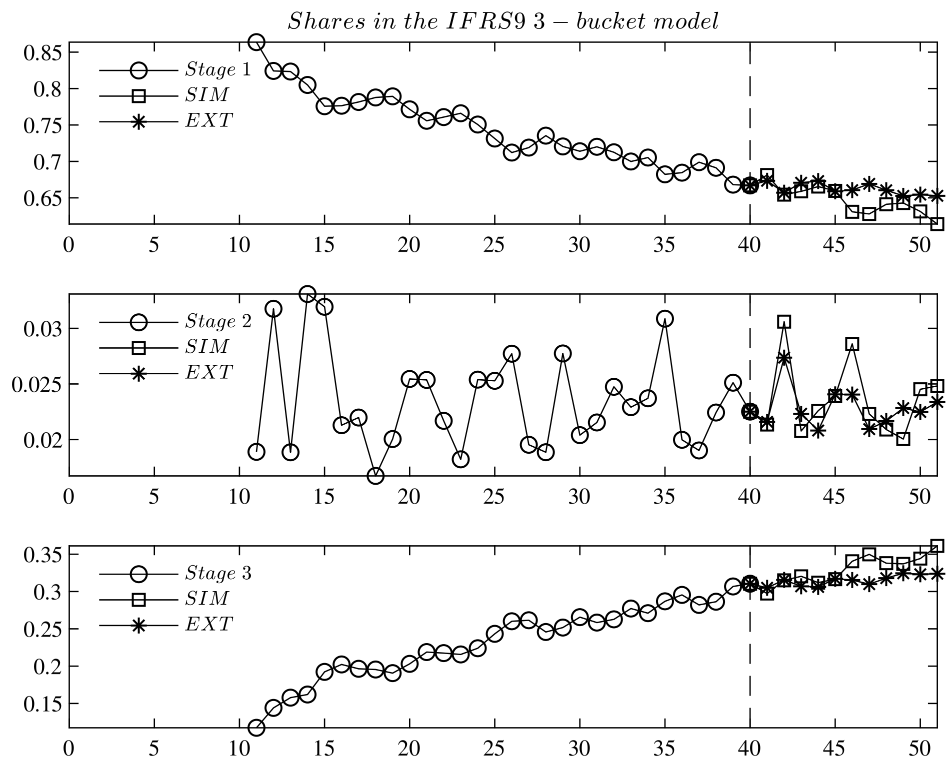

$Entries\ at\ t:\ E_t$
$Stayers\ at\ t:\ S_t$
$Portfolio\ at\ t:\ F_t$
$Leavers\ at\ t:\ L_t^d$



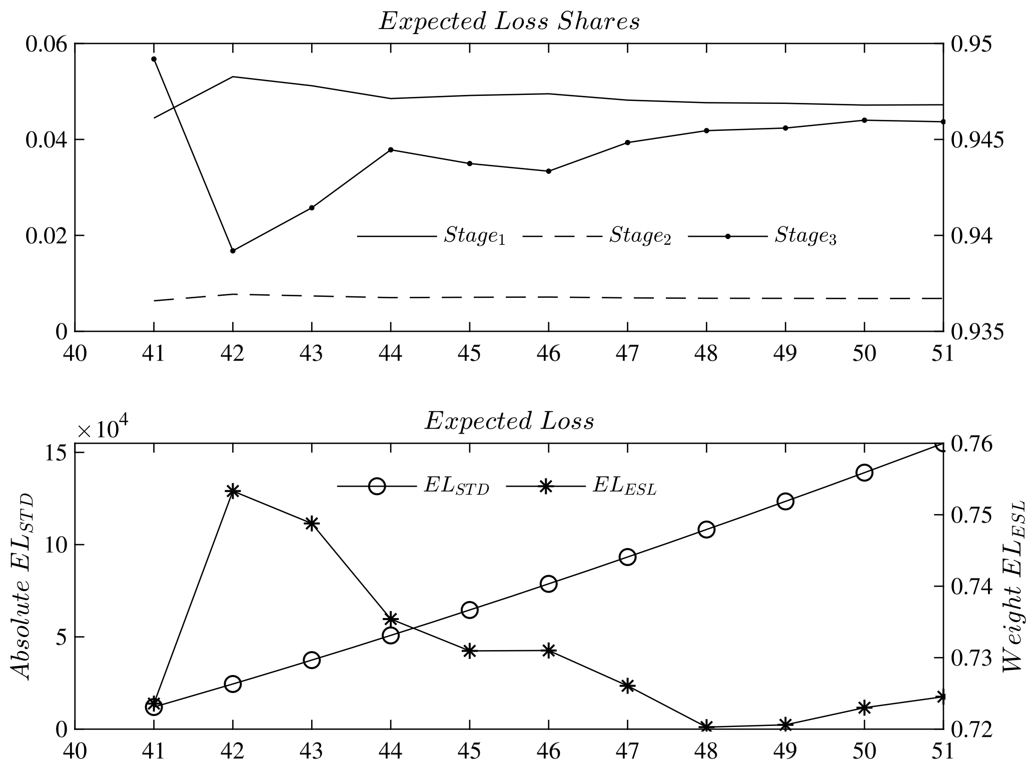
A_t
$\times 10^4$
1.6
1.4
1.2
1
0.8
15 20 25 30 35 40 45 50
B_t
12000
10000
8000
6000
4000
15 20 25 30 35 40 45 50
C_t
2400
2200
2000
1800
1600
1400
1200
1000
15 20 25 30 35 40 45 50
D_t
14000
12000
10000
8000
6000
4000
2000
15 20 25 30 35 40 45 50

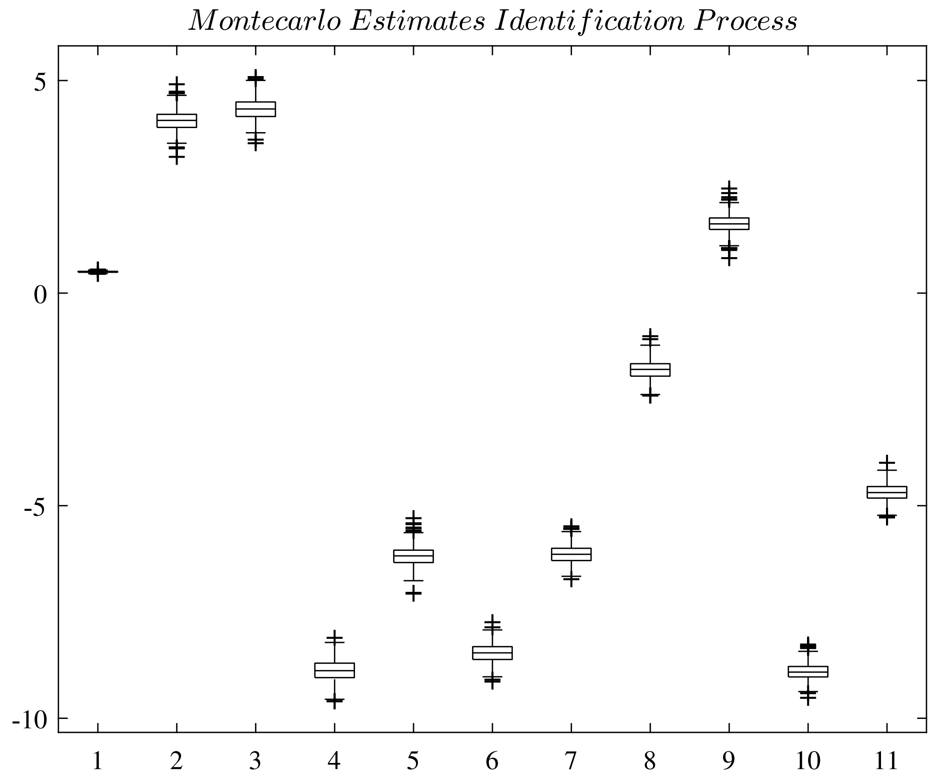


Absolute values in the IFRS9 3 − bucket model
×10^4
Stage 1
SIM
EXT
Stage 2
SIM
EXT
Stage 3
SIM
EXT

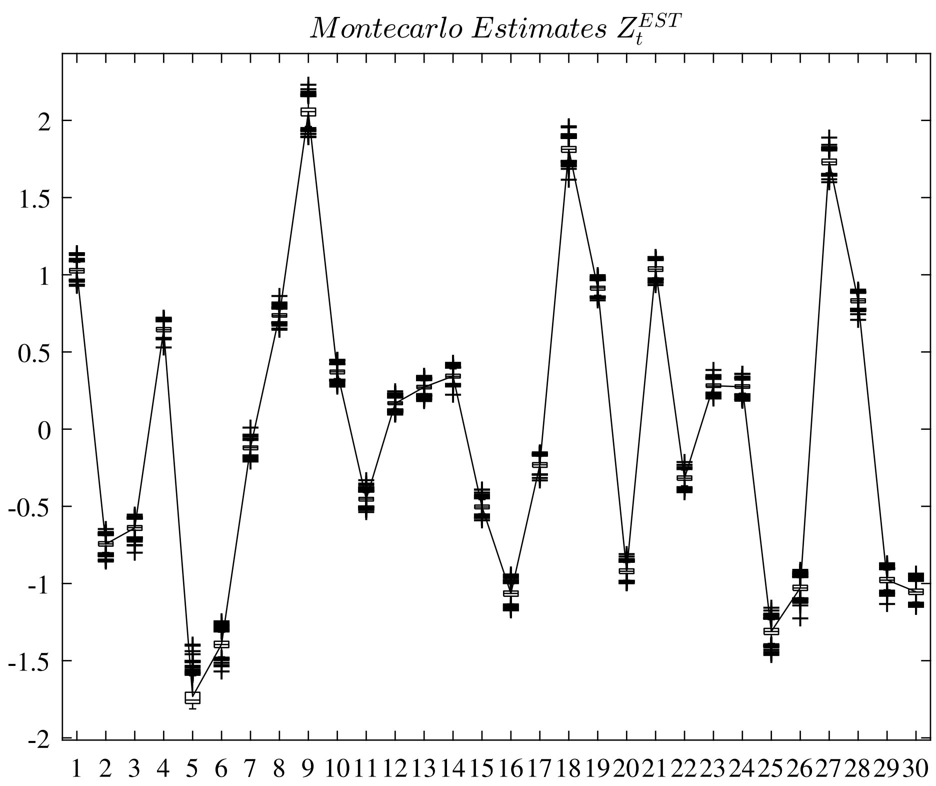


$Shares\ in\ the\ IFRS9\ 3-bucket\ model$
Stage 1
SIM
EXT
Stage 2
SIM
EXT
Stage 3
SIM
EXT

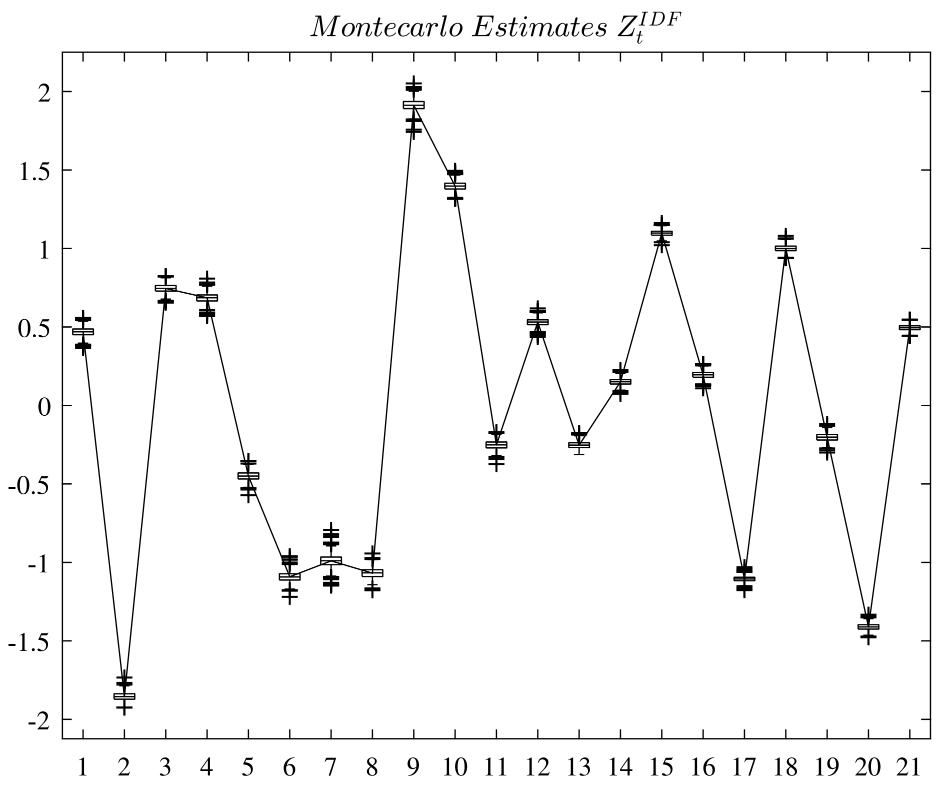

Expected Loss Shares
Stage1
Stage2
Stage3
0.06
0.04
0.02
0
0.95
0.945
0.94
0.935
40
41
42
43
44
45
46
47
48
49
50
51
Expected Loss
×10^4
Absolute EL_STD
Weight EL_ESL
EL_STD
EL_ESL
15
10
5
0
0.76
0.75
0.74
0.73
0.72



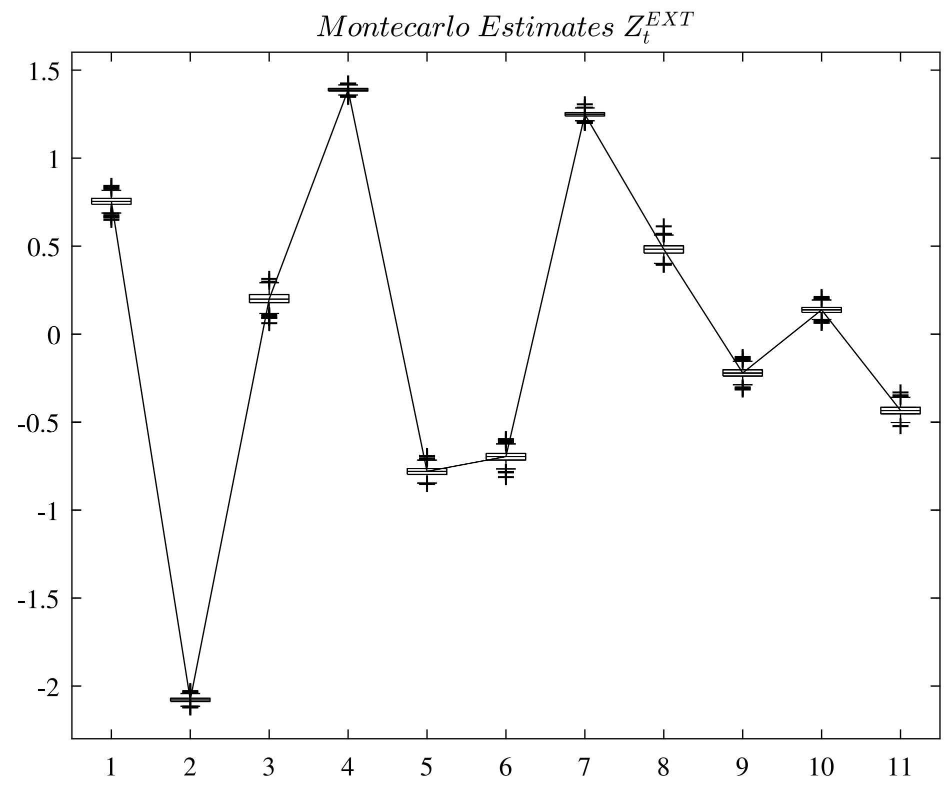
Montecarlo Estimates Identification Process
5
0
-5
-10
1
2
3
4
5
6
7
8
9
10
11

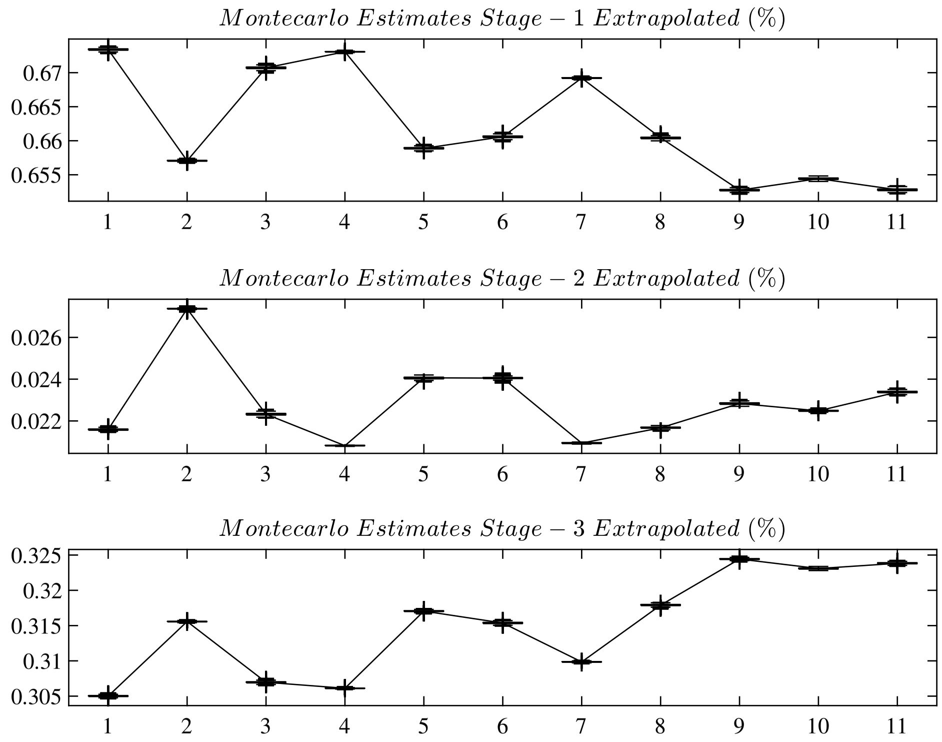


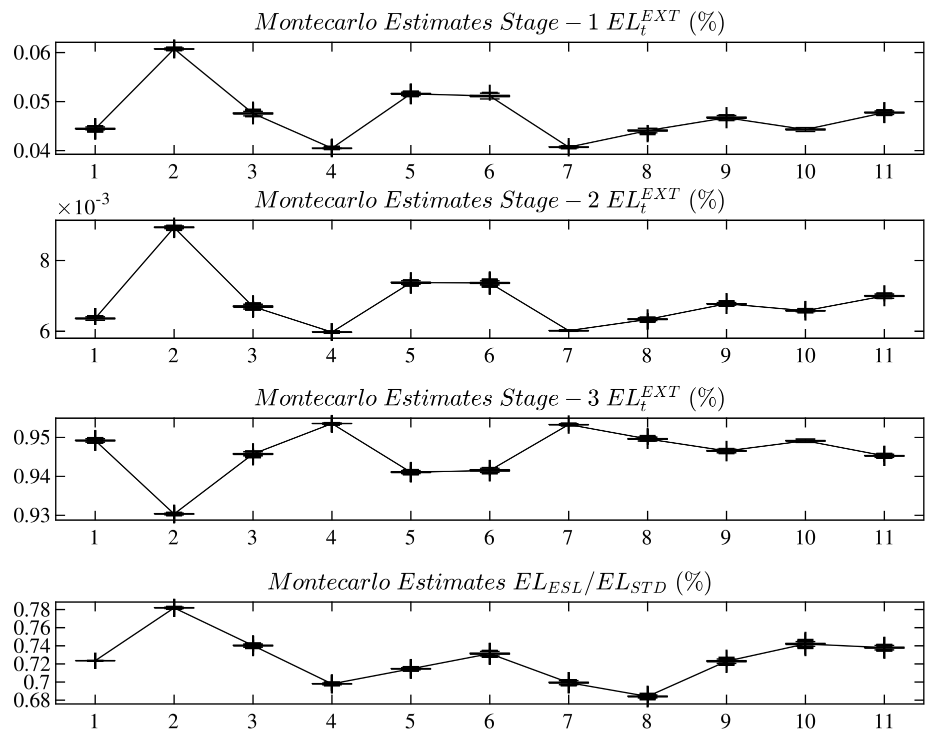
$Montecarlo\ Estimates\ Z_t^{EST}$
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30



$Montecarlo\ Estimates\ Z_t^{IDF}$
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21


$Montecarlo\ Estimates\ Z_t^{EXT}$
1.5
1
0.5
0
-0.5
-1
-1.5
-2
1
2
3
4
5
6
7
8
9
10
11


$Montecarlo\ Estimates\ Stage-1\ Extrapolated\ (\%)$
0.67
0.665
0.66
0.655
1 2 3 4 5 6 7 8 9 10 11
$Montecarlo\ Estimates\ Stage-2\ Extrapolated\ (\%)$
0.026
0.024
0.022
1 2 3 4 5 6 7 8 9 10 11
$Montecarlo\ Estimates\ Stage-3\ Extrapolated\ (\%)$
0.325
0.32
0.315
0.31
0.305
1 2 3 4 5 6 7 8 9 10 11


$Montecarlo\ Estimates\ Stage-1\ EL_t^{EXT}$ (%)
0.06
0.05
0.04
1 2 3 4 5 6 7 8 9 10 11
$\times 10^{-3}$
$Montecarlo\ Estimates\ Stage-2\ EL_t^{EXT}$ (%)
8
6
1 2 3 4 5 6 7 8 9 10 11
$Montecarlo\ Estimates\ Stage-3\ EL_t^{EXT}$ (%)
0.95
0.94
0.93
1 2 3 4 5 6 7 8 9 10 11
$Montecarlo\ Estimates\ EL_{ESL}/EL_{STD}$ (%)
0.78
0.76
0.74
0.72
0.7
0.68
1 2 3 4 5 6 7 8 9 10 11