Value Proposition

Open Source Financial Predictive Analytics

union
legend

February 2009

http://www.union-legend.com/ecap.pdf
http://www.revolution-computing.com/industry/finance.php

Retrospect and Currency

- The source of the Credit Crisis was in large part crucial failures in internal reporting and IT systems which comply with "Transparency Standards".

- Bradford & Bingley, Northern Rock, UBS, Credit Suisse; all referred this explicitly, in the public domain as did Mervyn King (BoE Governor).

- The Transparency Standards are being toughened by the Governments worldwide, right now; to be defined finally after the G20 in London in April 2009.

- Reviews of regulation in US and UK are coordinated through the FSF (Financial Stability Forum) driven by Gordon Brown.

- There is general consensus that the banks and insurance companies never even met the transparency standards in place in 2006. Now they are being 'enhanced', 'tightened'.

Open Source Financial Predictive Analytics

Innovation is the route to exit recession

- There is no question that in general black box proprietary closed source predictive analytics have failed the banking industry and thus society.

- On the other hand the large scale data management platforms for banking from IBM and SAP cannot be dismissed nor bettered. They have invested so much intellectual capital in these platforms it would take millennia for Open Source to catch up in that layer of the stack.

- The next step has to be about Open Source, almost certainly with a commercial backing in terms of support. Open Source is not exclusively about Predictive Analytics, it's just that the Community aspect is eminently applicable to predictive analytics, since the problems are generally hard and are generally iteratively solved.

- So at the top layer it is commercialized Open Source, allowing the community and forge to iteratively investigate and refine the quantitative analytics we need to understand our highly complex world.

- The answer to the current predicament is Open Source Financial Predictive Analytics as the top layer of a stack predicated on the banking data management platforms of IBM and SAP.

- The Central Banks, Universities, Software Vendors, individual developers and consulting firms are constantly publishing papers in the public domain about how to do modern risk management, most of these model risk management in R.

Community Open Source

Banking Supervision / Risk Quantification

- The Community Open Source idea in relation to meeting Banking Supervision requirements is just the kind of innovation in thinking and technology that Europe should be exhibiting in the middle of this recession, to develop a path to the future and exit this recession on a stable trajectory such that a Crisis like the one we are in will never happen again.

- The detailed evidenced argument that the manner of implementing risk analytics is already done by the Central Banks and Academics is out there, this entails that there is no need for any single financial entity to re-invent the whole domain intellectual capital of macroeconomic stress testing.

- Most of the macroeconomic stress testing has been conducted in R by the Central Banks and Universities.

- Through Community sharing, statisticians and economists in the banks can begin the process of development of macroeconomic through the cycle risk and capital quantification.

- Only through the cycle (a priori not pro cyclical) risk capital estimation predicated upon either full blown DSGE models or pragmatic stochastic positivism (Professor McNeil) will enable the reopening of the wholesale credit markets in Europe.

Solving the analytic and transparency bit

The Argument for R

- Commercial econometric software in the US started in Boston at the Massachusetts Institute of Technology (MIT), more specifically at the Center for Computational Research in Economics and Management Science. In the 1970s, code developed at MIT was not really copyright protected; it was built to be shared with the FED and other universities.

- Through the 60s and 70s various statistical modeling packages for economics were built particularly at Wharton, the University of Michigan and the University of Chicago (where the Cowles Commission had been located). At Princeton the focus was on development of econometric models in FORTRAN. The use of FORTRAN is much declining now but Chris Sims, now at Princeton, who developed the VAR methodology in an applied manner and was at the forefront of RE in the 1970s now makes all his models freely available in R.

- More and more econometricians are switching to the freely-available statistical system R. Free procedure libraries are available for R, http://www.r-project.org , an Open Source statistical system which was initiated by statisticians Ross Ihaka and Robert Gentleman.
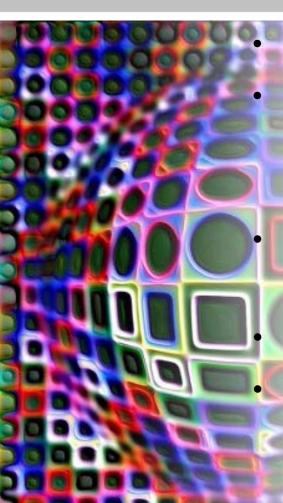
Banking Supervision

- Non-Linearities, Model Uncertainty, and Macro Stress Testing, by Miroslav Misina and David Tessier, Working Paper/Document de travail, 2008-30; Banque de Canada.
- Macro-model-based stress testing of Basel II capital requirements, Esa Jokivuolle, Kimmo Virolainen & Oskari Vähämaa, Bank of Finland Research, Discussion Papers, 17 2008.
- SIMULATING FINANCIAL INSTABILITY, Conference on stress testing and financial crisis simulation exercises, The European Central Bank, July 2008.
- Stress testing of real credit portfolios, Ferdinand Mager, Christian Schmieder, Discussion Paper, Series 2: Banking and Financial Studies, No 17/2008.
- Modelling The Distribution Of Credit Losses With Observable And Latent Factors, Gabriel Jiménez and Javier Mencía, 2007, Documentos de Trabajo, No. 0709, Banco de Espana.
- Stress Tests for the Austrian FSAP Update, 2007: Methodology, Scenarios and Results, Michael Boss, Gerhard Fenz, Gerald Krenn, Johannes Pann, Claus Puhr, Thomas Scheiber, Stefan W. Schmitz, Martin Schneider and Eva Ubl,
- The Next Generation of Default Prediction Models, Andreas Blochlingery, Zurcher Kantonalbank, Version: April 2007.
- A FRAMEWORK FOR STRESS TESTING BANKS' CREDIT RISK, Research Memorandum 15/2006, October 2006, Hong Kong Monetary Authority.
- The Basel II framework: the role and implementation of Pillar 2, PIERRE-YVES HORAVAL General Secretariat of the Commission Bancaire, Banque de France Financial Stability Review No. 9 December 2006.
- DEVELOPING A FRAMEWORK FOR STRESS TESTING OF FINANCIAL STABILITY RISKS, NIGEL JENKINSON, Executive Director, Financial Stability, Bank of England, 2007.
- Integrating credit and interest rate risk: A theoretical framework and an application to banks' balance sheets Mathias Drehmann, Steffen Sorensen & Marco Stringa; First Draft, April 2006.
- Adjusting Multi-Factor Models for Basel II-consistent Economic Capital by Marc Gürtler, Martin Hibbeln, and Clemens Vöhringer, summer 2008.
- A Tractable Model to Measure Sector Concentration Risk in Credit Portfolios, Klaus Düllmann and Nancy Masschelein, version: October 2006.
- Stress testing as a tool for assessing systemic risks Bank of England Financial Stability Review: June 2005.
- Liquidity Stress-Tester: A macro model for stress-testing banks' liquidity risk, Jan Willem van den End, Netherlands Central Bank, Research Department in its series DNB Working Papers 175, May 2008.
- Credit Risk Factor Modeling and the Basel II IRB Approach, Alfred Hamerle, Thilo Liebig, Daniel Rösch, Deutsche Bundesbank, Preliminary Draft from: October 7, 2002.
- Stresstests in Banken, Von Basel II bis ICAAP, Kai-Oliver Klauck, Claus Stegmann, iFB, 2006.

http://www.union-legend.com/ecap.pdf
http://www.revolution-computing.com/industry/finance.php

6

Financial Technology

Information Technology

- The mathematicians and econometricians are going to take over the bank, transparency in complexity is demanding that now!

- All this spin blaming mathematics for the credit crisis is just spin from middle management who never understood it anyway, protecting their reputations, it was their failure to put quantitative analytics at the top of the systems and thinking agenda which caused this crises and which is stretching it out longer than is necessary.

- We are systemically dependent upon innovations in financial technology now. Computation of risk capital in an holistic and comprehensive manner is the key to recovery from this crisis episode.

- **We have to meet the complexity of our financial technology needs with our response in terms of information technology!**

- The key business accelerator in Open Source is Community, particularly in Financial Predictive Analytics; since it is via the community in an open source framework that one's initial intellectual capital is gathered.

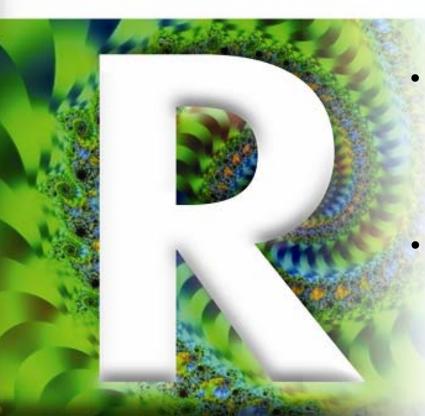# High Performance Computing (HPC)

The Challenge

- HPC for R and support for the technology are necessary to set open source predictive analytics to work in a "production" environment.

- In some instances the predictive techniques are so computing intensive that they require to run for 10s of hours on a standard, even large scale system rack, that is why predictive analytics draws in the requirement for the High Performance Computing (HPC) platform, which means putting clusters of processers together in parallel to run simultaneously at one problem rather than running sequentially at several problems.

- Good data modeling cannot help time series data, time series just is what it is.

- R is designed by statisticians for statisticians. The implication is that it is not designed by computer scientists, let alone HPC experts.

- It is crucial to retain the attention and "affection" of users, (stickiness in a way) to a complex development environment like that necessary for econometric modeling

- R runs in serial if left to its own devices. On datasets where really compute intensive models are applied it can simply run for hours or run out of memory/crash or both.

- In some instances the predictive techniques are so computing intensive that they require to run for 10s of hours on a standard, even large scale system rack.

High Performance

- REvolution enhance R to make it scale in parallel. There is scale up parallelism (which takes advantage of high performance numerics optimized for Intel x86 architectures that REvo has privileged access to) and scale out parallelism (using REvo's parallelR) to run models on arbitrarily large clusters.

- REvolution compile R within REvolution R Enterprise (to GAMP 5 standards) and integrate both scaling techniques, so that the R user has no additional or different programming to do. It "just works", and is a supported platform for research, production and regulated environments.

- Financial Predictive Analytics draws in the requirement for the High Performance Computing (HPC) platform, which means putting clusters of processers together in parallel to run simultaneously at one problem rather than running sequentially at several problems.

Support

Commercial Open Source

- HPC for R and support for the technology are necessary to set open source predictive analytics to work in a "production" environment.
- R is designed by statisticians for statisticians. The implication is that it is not designed by computer scientists, let alone HPC experts.
- Many times a user can get confused between what is a code issue and a methodology issue.
- The User of course is not like any 'normal' user, the user is already probably a PhD or at least Masters educated.
- There are commercial challenges to managing an Open Source project in-house. Some Banks can do it but it is "new stuff". All too often we ignore the real effects of "new stuff" on our project timescales. The impacts of technology change on IT projects in general and on the planning process in particular is often underestimated. You need to have the right knowledge and experience about the new stuff; methodology and development techniques; to plan and execute on its implementation.
- There is no way that an ordinary user can manage all of the disparate communities and forges which exist and may develop in an open source development language that is they key deliverable of the support centre to appraise, unit test and understand all of the objects out there in the social or community networks.

REvolution Computing

Support for R

- The commercial open source vendor provides a documented, supported build together with recourse in the event of issues with the software, just like any proprietary software vendor. REvolution Computing is the same in its relationship to R in all these respects. Enterprise Mission Critical needs have very real concerns behind them and REvolution Enterprise-level customer support makes REvolution products suitable for professional, commercial and regulated environments. REvolution provides technical support from statisticians and computer scientists highly versed in the R language and the specifics of each REvolution R build; http://www.revolution-computing.com/support/

- That is what REvolution Computing provides for R, the Open Source development community environment in which all of the intellectual capital in econometrics since 1946 is embedded. The open source community of worldwide statisticians and econometricians is at the bleeding edge of analytics, but not always creating software that can be set to work in a scale-able fashion, from an IT or production software perspective (it is not designed with that in mind). REvolution is aimed at ensuring that it can – i.e. ensuring that the best from the research world can be used in production.

REvolution
computing

We do the math

Union Legend - Asymptotix

Financial Predictive Analytic Solution Architecture

- **The UL methodology** is to combine expert consulting with experienced resourcing. Our Consultants partner with your staff and industry experts so that your project is fully equipped with optimized human capital, providing the highest level of business and technical knowledge. This focused perspective brings project resource to our clients, which has been rigorously validated as appropriate to the specific requirement. UL can through this philosophy bring optimized specialist resource to bear on each complex task requirement within the overall framework of your project or programme plan.

- UL can assist our clients in the strategy, solution design and implementation of integrated Data Management and Financial Predictive Analytic (FPA) platforms. In other words, the UL mission in FPA is the integration of 3rd generation BI (Predictive) environments with standard BI toolsets and the data management platforms normally implemented to support them. We can further assist with the planning and architecture for High Performance Computing (HPC) environments for Predictive Analytics, where these may be necessary (and in our view they will be). UL has further relevant experience with the integration of General Ledger platforms with Risk Management data warehouses or marts and we have the specialist knowledge to integrate FPA tools into the GL environment. In short we can provide the support necessary to assist management to move towards Target Operating Models predicated upon the integration of FPA with the GL and operational BI for Credit, Market and Liquidity Risk analytics, necessary post CC.

- **UL is soon to be re-branded Asymptotix.**