Research Statement

Ba M. Chu*

My research focuses on important areas of financial econometrics, such as risk management, asset allocation and modelling dependence in finance, and of theoretical econometrics, such as time series analysis and non/semi-parametric inference for weakly dependent data. In the area of asset allocation, I develop a large-deviations approach to approximate the probability that a portfolio return underperforms a predetermined benchmark for well-diversified portfolios of financial assets. In the area of dependence modelling, I propose a maximum-entropy approach to recover relative entropy measures of dependence from limited information by constructing the most entropic copulas (MECs), which are then implemented in various optimal investment problems. It is empirically demonstrated that economic and statistical gains as a direct consequence of applying this approach are significant. In the area of time series analysis, I develop new asymptotic methods, such as central limit theorems and concepts of weak dependence, to study the behaviour of a class of semiparametric estimators in the large samples. The next few sections describes the major projects that I have been working on.

1 Optimal Investment Problems with Large Portfolios

Hedge funds and pension funds often face the problem of choosing optimal asset allocation strategies for portfolios invested in a large number of investable stocks. This research programme focuses on developing a new method based on the large deviations theory to choose an optimal portfolio such that the risk, which is defined as the probability that the portfolio return underperforms

*Department of Economics, Carleton University, B-857 Loeb Building, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada. Phone: +1 (613) 520 2600 (ext. 1546). Fax: +1 (613) 520 3906. E-mail: ba_chu@carleton.ca. Web Page: http://http-server.carleton.ca/~bchu/.

an investable benchmark, is minimal. As a particular case, the proposed study also analyzes the effect of two types of asymmetric dependence; 1) asymmetry in a portfolio return distribution and 2) asymmetric dependence between asset returns; on the optimal portfolios invested in two risky assets. Calibration of the proposed method to real data confirms that the asset allocation strategy suggested by the theory is indeed optimal as one holds a large portfolio.

To give some further details of the proposed approach, the paper (P1) "Large Deviations Theorems for Optimal Investment Problems with Large Portfolios" (joint with John Knight and Stephen Satchell) [published in the *European Journal of Operational Research*] proposes a new theoretical framework, based on *large deviations theory*, for the problem of optimal asset allocation in large portfolios. This problem is, apart from being theoretically interesting, also of practical relevance; examples include, *inter alia*, hedge funds where optimal strategies involve a large number of assets. In particular, we also prove the upper bound of the shortfall probability (*or* the risk bound) for the case where there is a finite number of assets. In the two-assets scenario, the effects of two types of asymmetries (i.e., asymmetry in the portfolio return distribution and asymmetric dependence among assets) on optimal portfolios and risk bounds are investigated.

In the paper (P2) "Large Deviations Estimation of the Windfall and Shortfall Probabilities for Optimal Diversified Portfolios [forthcoming in the *Annals of Finance*], I develop asymptotic approximations of the *windfall* and *shortfall* probabilities for an optimal portfolio of risky assets pertaining to the one-period investment problem as the number of the assets becomes sufficiently large. This is a practically interesting problem for the following two reasons: first, as having explained from the outset, investors can, for most of the time, reduce idiosyncratic risks by holding diversified portfolios. Second, diversification is more relevant (and indeed less costly) for short-term investments than long-term ones. In the long term, diversification is not necessary for the reason that risky assets appear safer over a longer time frame as idiosyncratic risks decline over time. And diminishing idiosyncratic risks are due to the fact that more information about firms become available at investors' disposal – that is, the longer one times the horizon the more predictable returns become. Thus, reminiscent of the results of Barberis (2000), in a portfolio containing cash and a stock, the allocation to the stock increases with the horizon.

2 Dependence Modelling with Copulas

Modeling dependence between asset returns is crucial in a number of financial applications. For a number of standard distributions in the elliptical family, dependence is simply measured by Pearson's correlation coefficient. In practice, however, asset returns do not always belong to the elliptical family, and thus the dependence structure does not always show out of the joint distribution function under consideration. It would be useful to separate the statistical properties of each return from their dependence structure. Copulas provide us with a viable way to achieve this goal.

The paper (P3) "Recovering Copulas from Limited Information and An Application to Asset Allocation [published in the *Journal of Banking and Finance*] proposes an *entropy*-based method to construct a new class of copulas – the most entropic canonical copulas (MECC). The empirical study focuses on an investment problem for an investor with a constant relative risk aversion (CRRA) utility function allocating wealth between the Dow Jones Large-Cap and Small-Cap indices, of which the *contemporaneous dependence* can be modeled by the MECC or other commonly-used copulas. Both the theoretical analysis of the method and the empirical study indicate the potential for enormous statistical and economic gains as a result of using the MECC. The theoretical justification of the MECC as a typically robust class of copulas is provided in the paper (P4) "Most Entropic Copulas (joint with S. Satchell).

3 Time Series Econometrics

Central limit theorems (CLTs) and functional central limit theorems (FCLTs) for the summands of weakly dependent processes are indispensable tools to derive limiting distributions for many econometric estimators and test statistics.

In the paper (P4) "Limit Theorems for the Discount Sums of Moving Averages" [forthcoming in Journal of Time Series Analysis], I establish standard conditions so that the discount sum of MAs – defined as $S_T(\phi) = \sum_0^T \phi^t \widetilde{\xi}_t$, where $\widetilde{\xi}_t = \mathcal{C}_0(L)\xi_t = \sum_{i=0}^{\infty} a_i \xi_{t-i}$, in which $\mathcal{C}_0(L)$ is a lag polynomial, and $\phi \in (0, 1)$ – converges to normality when ϕ approaches 1 from below, thus effectively extends the result by Omey (1984) into the non-i.i.d. case. Proofs of the results obtained in this paper utilize the BN decomposition and some standard assumptions that make the higher-order terms in the BN decomposition vanish asymptotically. As a matter of fact, the discount sum is important in certain empirical applications. One example is the equity pricing model in corporate finance. This model asserts that the current fair price, P_0, of a stock is equal to the sum of all the future dividends, D_t $\forall t = 1, \ldots, T$, discounted back to the current time. That is, $P_0 = V_T$, where $V_T = \sum_{t=1}^{T} \phi^t D_t$, and ϕ^t is a discount factor (for instance, $\phi = \frac{1}{1+r}$, where r is an interest rate). Hence, P_0 is determined by the behavior of V_T. Nonetheless, the limiting behavior of V_T can hardly be inferred, even though V_T is a convergent series. Assuming that the dynamic of D_t is a MA and that ϕ is close to one from below, then V_T is the discount sum of MAs. It is also noteworthy at this point that, in reality, given a normal market condition the interest rate, r, is often so moderate that the discount rate, $\frac{1}{1+r}$, becomes rather close to one. In this case, one could anticipate that the sample distribution of V_T can be legitimately approximated by the limiting distributions stated in Theorems **1** and **2** of the paper.

In the paper (P5) "Testing Distributional Assumptions: A L-Moment Approach" (joint with Mark Salmon), we propose a test for conditional parametric distribution functions with weakly dependent stationary time-series data. The test is based on moment conditions which can be uniquely determined by Stein's equations and the L-statistics of conceptual ordered subsamples drawn from the population sample of a distribution; hereafter referred to as the generalized method of L-moment (GMLM) test. Test limiting distributions are χ-squared; the covariance kernels in the test reflect parametric dependence specification and parameter estimation error. The GMLM test can resolve the choice of orthogonal polynomials remaining as an identification issue in the GMM tests using the Stein approximation (Bontemps and Meddahi; 2005, 2006). In addition, since L-moments can be represented as the moments of the quantiles of a distribution, the test functions in the Stein approach can be represented as the functions of the quantiles of the conditional distribution. In addition, an application to elliptical conditional distribution for autoregressive processes is provided.

Furthermore, we also discuss the issue of robustness against dynamic misspecification. Finally, we provide some Monte-Carlo simulation to examine the size, the power and the robustness of the GMLM test and compare it with both existing parametric and bootstrap conditional distribution tests.

4 Nonparametric/Semiparametric Econometrics

Functionals of order statistics and their concomitants, and the k-nearest neighbour (NN) method provide viable alternatives to classical nonparametric methods using Nadaraya-Watson kernels.

In the paper (P6) "Linear Functionals of Concomitants of Order Statistics with Dependent Data, (joint with D. T. Jacho- Chávez), we establishes the $\sqrt{T}$-asymptotic normality of statistics of the form $T^{-1}\sum_{t=1}^{T} J(t/T)h(X_{(t)}, Y_{[t]})$, where J is a bounded smooth weighting function, h is some real-valued function of (x, y), $X_{(t)}$ represents the t-th order statistics and $Y_{[t]}$ its concomitant from the strictly stationary and ergodic data $\{X_t, Y_t\}_{t=1}^{T}$. The utility of these results are illustrated by establishing the asymptotic normality of Yang's (1981) estimator of the regression function $E[Y|X = x]$ under strong-mixing conditions. An interesting point to note about this estimator of $E[Y|X = x]$ is that 1) the estimator contains no random denominator, thus, unlike the kernel regression, trimming is not needed; 2) in terms of the first-order asymptotics, this estimator is not, in the limit, perturbed by any degree of weak serial dependence in the data – this property is therefore desirable for nonparametric bootstrap procedures using the estimator because an i.i.d. bootstrap algorithm could then be used instead of rather complicated block bootstraps procedures.

In the paper (P7) "k-Nearest Neighbour Estimation of Inverse-Density-Weighted Expectations with Dependent Data, (joint with D. T. Jacho- Chávez) [forthcoming in *Econometric Theory*], we study the problem of estimating expected values of functions that are inversely weighted by an unknown density using the k-Nearest Neighbour (k-NN) method. It establishes the $\sqrt{T}$-consistency and the asymptotic normality of an estimator that allows for strictly stationary time-series data. Under the independent random sampling scheme the proposed estimator is also shown to be asymptotically semiparametric efficient. The consistency of the Barlett estimator of the asymptotic variance is also established. Monte Carlo experiments show that the proposed estimator performs well in finite sample applications.

5 Research Assessment

There are many complexities and subjectivities associated with evaluating one's own work, this section makes use of the 2010 edition of the *Journal Citation Report*® (JCR) by the Institute of Scientific Information (ISI), a division of Thomson Reuters. The JCR provides systematic and objective information on citation data in academic journals related to the sciences and social sciences.

Table 1 shows the following information: The **Impact Factor** (IF) of a journal is defined as the average number of citations in a year to those papers in a journal that were published during the 2 preceding years. The IF is frequently used as a proxy for the relative importance of a journal within its *main* and *sub*-fields. The **5-Year Impact Factor** (5-Year IF) is similar to the IF, however calculations are performed using the last 5 preceding years instead. The article **Influence Score**™ (IS) measures how often articles in the journal are cited within the first five years after its publication. Finally the table also shows the overall **ranking** of each journal based on their respective IF in their main category (Rank 1) and sub-category (Rank 2) as classified by the 2010 JCR.

Remark 5.1. Annals of Finance provides an outlet for original research in all areas of finance and its applications to other disciplines having a clear and substantive link to the general theme of finance. Unfortunately, it is not part of the ISI database, and it is therefore excluded from most rankings based on the JCR. However, leading researchers in my field such as Rustam Ibragimov, Johan Walden, Walter Schachermayer, and Jakša Cvitanić, among many others, have published their work in this journal.

Remark 5.2. Asia-Pacific Financial Markets is the official journal of the Japanese Association of Financial Econometrics and Engineering. Unfortunately, it is not part of the ISI database, and it is therefore excluded from most rankings based on the JCR. However, leading scholars in the fields of Mathematical Finance such as Jirô Akahori , Robert Elliott , Dilip Madan, and Eckhard Platen, among many others, have published their work in this journal.

Remark 5.3. Kalaitzidakis, Mamuneas and Stengos (2003) provide the economics journal ranking based on citations in 1998 of articles published only in 1994-1998 excluding self-citations and adjusted for impact (influence) and size. Kalaitzidakis, Mamuneas and Stengos (2003, Table 1) shows that Econometric Theory ranks number seven, whilst the top-field journal (Journal of Econometrics) ranks number six.

References

Barberis, N. (2000). Investing for the long run when returns are predictable, *Journal of Finance* **55**: 225–264.

Bontemps, C. and Meddahi, N. (2005). Testing normality: A gmm approach, *Journal of Econometrics* **124**: 149–186.

Bontemps, C. and Meddahi, N. (2006). Testing distributional assumptions: A GMM approach, *Working paper, Universite de Montreal* .

Kalaitzidakis, P., Mamuneas, T. P. and Stengos, T. (2003). Rankings of academic journals and institutions in economics, *Journal of the European Economic Association* **1**(6): 1346–1366.

Omey, E. (1984). A limit theorem for discount sums, *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **68**: 49–51.

Yang, S.-S. (1981). Linear functions of concomitants of order statistics with application to nonparametric estimation of a regression function, *Journal of the American Statistical Association* **76**(375): 658–662.

Table 1: 2008 JCR Science & Social Science Edition

Journal	IF	5-Year IF	IS	Rank 1	Rank 2
JOURNAL OF BANKING & FINANCE[a]	2.731	2.528	0.800	23/305	6/76
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH[b]	2.159	2.513	0.886	N/A[e]	6/75
STUDIES IN NONLINEAR DYNAMICS AND ECONOMETRICS[a]	0.765	0.968	0.599	148/305	23/43
OPTIMIZATION[c]	0.509	0.755	0.528	177/236	62/75
FINANCE RESEARCH LETTERS	0.314	N/A	N/A	N/A	66/76
JOURNAL OF TIME SERIES ANALYSIS[d]	0.678	0.888	0.871	77/110	64/93
ECONOMETRIC THEORY[a]	1.015	1.264	1.541	107/305	21/43

[a] The main field for Economics (Finance) Journals is 'Economics', whilst the sub-field is 'Mathematical Methods' ('Business and Finance').

[b] The main field is 'Operations Research & Management Science'.

[c] The main field is 'Applied Mathematics', whilst the sub-field is 'Operations Research & Management Science'.

[d] The main field is 'Statistics & Probability', whilst the sub-field is 'Mathematics, Interdisciplinary Applications'.

[e] 'N/A' is the abbreviation of 'Not Available'.