別表5
(3)

主　論　文　要　旨

No.1

報告番号	甲　乙　第　　号	氏　名	マクリン謙一郎

主 論 文 題 名：Essays on Bayesian Econometrics for Big Data

（内容の要旨）

This dissertation discusses Bayesian methods for econometric analysis of Big Data. Recent advancements in computational scalability and the increase in availability of large datasets has brought on a wave of interest in the analysis of large-scale data, with applications in many fields, from genetics, macroeconomic, finance, marketing, and so on. Here, "Big Data" is defined as datasets large enough where conventional statistical and econometric methods (such as ordinary least squares) fail to produce useful results that can be used for inference, forecasting, and decision making. While many economic data are, by nature, "big," much of the focus on econometrics has been on relatively small data. However, employing these tools for large-scale data has posed several problems, in particular, when model misspecification lead to biased results that can skew inference, forecasts, and decision making. In this regard, development of econometric methods (in terms of modeling, estimation, and computation) for these data is imperative. In this dissertation, several econometric and computational methods are developed to analyze large-scale data in finance and macroeconomics. These methods respond to specific problems and interests in the field, and provide important insight into the underlying questions using big data.

Chapter 1 introduces the ideas and methods discussed in this dissertation. Chapter 2 analyzes a large-scale cross-section of stock returns using a non-linear leverage stochastic volatility model. I analyze a plethora of stocks from the U.S. and Japan and find that, a well-known effect observed in analysis of stock indices, is not persistent in its original form for most stocks. Chapter 3 develops a parallel computational algorithm to estimate complex models in a fast on-line manner in order to analyze large-scale data. This responds to the need to analyze data sequentially in an efficient manner in order to analyze a large cross section of data. The algorithm developed uses the parallel computing architecture of graphical

processing units (GPUs) to achieve the speed needed for this type of analysis.

Chapter 4 develops a dynamic sparse factor model in order to analyze a large macroeconomic panel data. While factor models have proven to be a powerful tool for analyzing big data, it is not without its challenges. Here, I develop a novel method to analyze large panel data without the constraints of other methods, in a fast and efficient manner using the EM algorithm. Chapter 5 introduces a novel framework for large-scale predictive modeling and decision making. Modeling with a large number of covariates is a difficult, yet critical, problem in econometric analysis. However, available methodologies lack in terms of effective forecasting and decision making imperative for economists. I propose a framework that successfully models a large number of covariates, while retaining interpretability that is needed for decision making.