ESTIMATION FOR HIGH DIMENSIONAL DATA
IN MULTI-LEVEL MODEL

by

Ella Joyce S. Paragas

A thesis submitted in partial fulfillment
of the requirements for the degree

Master of Science (Statistics)

School of Statistics
University of the Philippines
Diliman, Quezon City

March 2014
ABSTRACT

Modeling of high dimensional data is often confounded with multicollinearity and problem with interpretability of the fitted model. General Adaptive Sparse Principal Component Analysis (GAS-PCA) is used in reducing dimensionality that simultaneously induces sparsity. However, selection of few sparse components of the high dimensional predictors leads to the specification bias. A random group level effect can help mitigate the bias in a model based on a few principal components.

Keywords: High Dimensional data, Multilevel model, General Adaptive Sparse Principal Component, Principal Component Regression