

Model Uncertainty & Model Averaging Techniques

Shahram Amini Moghadam

Dissertation submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Economics, Science

Christopher F. Parmeter, Co-Chair
Richard A. Ashley, Co-Chair
Richard D. Cothren
Kwok P. Tsang

July 16th, 2012
Blacksburg, Virginia

Keywords: Mallows, Jackknife, Bayesian, Frequentist, Model Averaging, Nonparametric
© Copyright 2012, Shahram Amini Moghadam

Model Uncertainty & Model Averaging Techniques

Shahram Amini Moghadam

ABSTRACT

The primary aim of this research is to shed more light on the issue of model uncertainty in applied econometrics in general and cross-country growth as well as happiness and well-being regressions in particular. Model uncertainty consists of three main types: *theory uncertainty*, focusing on which principal determinants of economic growth or happiness should be included in a model; *heterogeneity uncertainty*, relating to whether or not the parameters that describe growth or happiness are identical across countries; and *functional form uncertainty*, relating to which growth and well-being regressors enter the model linearly and which ones enter nonlinearly.

Model averaging methods including Bayesian model averaging and Frequentist model averaging are the main statistical tools that incorporate theory uncertainty into the estimation process. To address functional form uncertainty, a variety of techniques have been proposed in the literature. One suggestion, for example, involves adding regressors that are nonlinear functions of the initial set of theory-based regressors or adding regressors whose values are zero below some threshold and non-zero above that threshold. In recent years, however, there has been a rising interest in using nonparametric framework to address nonlinearities in growth and happiness regressions.

The goal of this research is twofold. First, while Bayesian approaches are dominant methods used in economic empirics to average over the model space, I take a fresh look into Frequentist model averaging techniques and propose statistical routines that computationally ease the implementation of these methods. I provide empirical examples showing that Frequentist estimators can compete with their Bayesian peers. The second objective is to use recently-developed nonparametric techniques to overcome the issue of functional form uncertainty while analyzing the variance of distribution of per capita income. Nonparametric paradigm allows for addressing nonlinearities in growth and well-being regressions by relaxing both the functional form assumptions and traditional assumptions on the structure of error terms.

Acknowledgments

First, I would like to thank my adviser, Dr. Chris Parmeter. Chris, I am deeply grateful to you for teaching me everything I needed to know to accomplish this project. I have gained a lot of professional insight while working as your graduate student. Thanks for allowing me to vent when I needed it, and thanks for the motivation over the last four years. I could not have completed this research without your advice, patience, help, and friendship. You are the one who changed my life and made me believe in myself. Every single moment I have spent in your classes, in your office, in meetings, with your lovely family, etc. has been full of joy and happiness. I am indebted to you for that. Thank you again beloved Chris.

I would also like to independently acknowledge Dr. Richard Ashley. He not only provided helpful guidance as my thesis committee co-chair, but his instruction and mentorship throughout the different stages of my graduate education were integral to my development as a future scholar. He always made himself available to me when I needed to talk about my concerns and my needs. Dr. Ashley, I never forget your support and help. Without your succor, and encouragement this thesis would not have come to fruition.

I also have a great appreciation for Dr. Richard Cothren, Dr. Djavad Salehi-Isfahani, and Dr. Kwok Ping Tsang for their guidance as other members of my committee. Dr. Chotren's advanced macroeconomics class was the first class I attended here at Virginia Tech as a graduate student. I do not forget how stressed out I was before the class but he made it an interesting class with his sense of humor. Without Dr. Salehi's support during the process of selecting Ph.D. candidates, I would not have a chance to be in where I am now in my life. I still have kept the first email he sent me and asked me if I was interested in pursuing my education at the Department of Economics. I read his email many many times and did not feel my feet on the ground. How motivated and hard-working Dr. Tsang is has helped me think about the type of scholar I want to work to become. He is a young, humble professor who tirelessly works to succeed in his career.

Many thanks go to the faculty and staff of the Economics Department for valuable learning experiences and support over the past five years. I owe special thank to my entire family for their love, support, patience, and helping me become the person that I am today. They have always been there for me while I have confronted the ups and downs of life. I love you guys.

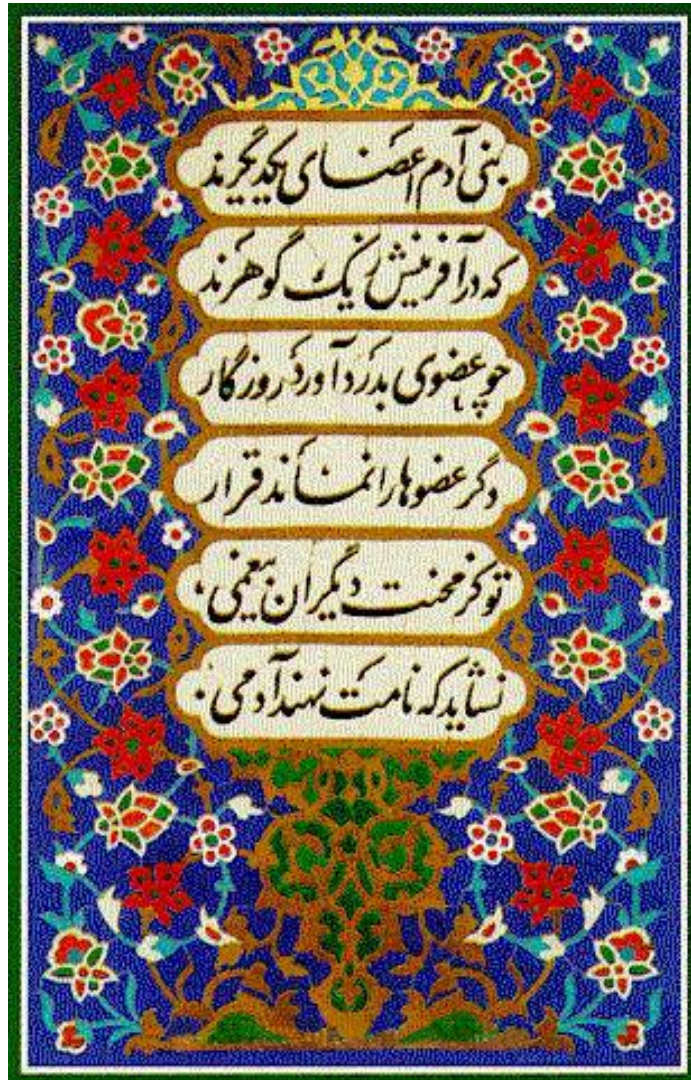
Finally, I would like to thank the people that made Virginia Tech a special and unique place for me. I met so many wonderful people here at Virginia Tech. To say I learned a great deal from them would be an understatement. Thank you Dear Grazia for your friendly advice, for listening to my stories, and for teaching me Italian language. Thank you Angela for being a great friend and thank you Nima for pizza nights!

My time at Virginia Tech has been full of hope and passion. With this degree completed, I move on to begin a career in economics or finance, which I have been eagerly awaiting. I offer my thanks to the rest of those who offered friendship and support along my journey.

Dedication

I would like to dedicate this thesis to my loving family and also those affected by the events of April 16, 2007 at Virginia Tech. Here is a poem by Saadi, a Persian poet, that I admire very much:

*Human beings are members of a whole,
In creation of one essence and soul.
If one member is afflicted with pain,
Other members uneasy will remain.
If you have no sympathy for human pain,
The name of human you cannot retain.*



Contents

1	Introduction to Model Averaging	1
1.1	Applied Economics & Model Uncertainty	1
1.2	Model Averaging Techniques	2
1.2.1	Bayesian Model Averaging	3
1.2.2	Frequentist Model Averaging	8
1.3	Nonparametric Statistics for Growth Empirics	12
2	Comparisons of Frequentist Model Averaging Techniques	15
2.1	Abstract	15
2.2	Introduction	16
2.3	Model Averaging Techniques	17
2.3.1	Bayesian Model Averaging	17
2.3.2	Weighted Average Least Squares	18
2.3.3	Frequentist Model Averaging	20
2.4	Data	23
2.5	Results	24
2.5.1	A comparison using Fernandez, Ley & Steel (2001 <i>b</i>)	25
2.5.2	A comparison using Masanjala & Papageorgiou (2008)	26
2.5.3	A comparison using Doppelhofer & Weeks (2009)	27
2.6	Jackknife Model Averaging	29
2.7	Jackknife Model Averaging Estimator	30
2.8	Model Space Orthogonalization	31
2.9	Data	33
2.10	Results	33
2.11	Conclusion	34
3	Weighting to be Happy: A Study of Determinants of Happiness	37

3.1	Abstract	37
3.2	Introduction	38
3.3	Model Uncertainty and Bayesian Model Averaging Analysis	39
3.4	Data	42
3.5	BMA results	44
3.6	Jointness of Happiness Determinants	45
3.7	BMA Analysis Across Countries	72
3.8	Conclusions	72
4	Conditional Variance Decomposition	88
4.1	Abstract	88
4.2	Motivation	89
4.3	Empirical Methods	92
4.3.1	Estimating a Conditional Density	93
4.3.2	Bandwidth Selection	94
4.3.3	Constructing the Variance	95
4.4	Empirics	96
4.4.1	Data	96
4.4.2	Results	97
4.5	Implications and Future Research	103
5	Conclusions, Contributions, and Future Work	105
	References	107
	Appendix A Bayesian Model Averaging in R	117
A.1	Abstract	117
A.2	Introduction	118
A.3	Available Packages for R	120
A.3.1	The BMS Package	120
A.3.2	The BAS Package	122

A.3.3	The BMA Package	124
A.4	Comparison of the Packages	126
A.4.1	Model Sampling	126
A.4.2	Model Priors	126
A.4.3	Alternative Zellner's g Priors	126
A.4.4	Outputs	128
A.4.5	Diagnostic Plots	128
A.5	Empirical Illustration	128
A.5.1	Enumeration of the Model Space	128
A.5.2	Model Sampling	129
A.5.3	Plot Diagnostics	132
A.6	Time Performance of the Three Packages	137
A.7	Replicability	138
A.8	Conclusions	142
Appendix B Computer Codes		143

List of Abbreviations

Δ	Data
Λ	Diagonal Matrix of Eigenvalues
ω_j	Weight Assigned to the j th Model
$\Phi(\cdot)$	Standard-Normal Cumulative Distribution Function
σ^2	Unknown Variance
$E(\cdot)$	Expected Value
$f(x, y)$	Joint Distribution
g	Zellner's g Prior
<i>i.i.d</i>	Independently Identically Distributed
k^c	Kernel Function for Continuous Variables
k^o	Kernel Function for Ordered Discrete Variables
k^u	Kernel Function for Unordered Discrete Variables
K_h	Product Kernel Function
N	Normal Distribution
Pr	Probability
Q	Orthogonal Matrix
AIC	Akaike Information Criterion
BIC	Bayesian Information Criterion
BMA	Bayesian Model Averaging
DW	Doppelhofer and Weeks
FIC	Focused Information Criterion
FLS	Fernandez, Ley, and Steel
FMA	Frequentist Model Averaging
GDP	Gross Domestic Product
HR	Harrison and Rubinfeld

JMA	Jackknife Model Averaging
LCLS	Local Constant Least Squares
LLLS	Local Linear Least Squares
MCMC	Markov Chain Monte Carlo
MENA	Middle East and North Africa
MISE	Mean Integrated Squared Errors
MMA	Mallows Model Averaging
MP	Masanjala and Papageorgiou
OECD	Organization for Economic Cooperation and Development
OLS	Ordinary Least Squares
PIP	Posterior Inclusion Probability
PMP	Posterior Model Probability
UIP	Unit Information Priors
WALS	Weighted Average Least Squares
BAS	Bayesian Adaptive Sampling
BMA	Bayesian Model Averaging
BMS	Bayesian Model Selection

List of Figures

1.1 A Visualization of Occam’s Window 5

A.1 Posterior Model Size Distribution and Model Probabilities Produced From
the **BMS** Package With “Uniform” Model Priors. 134

A.2 Posterior Model Size Distribution and Model Probabilities Produced From
the **BMS** Package With “Fixed” Model Priors. 135

A.3 Marginal Posterior Density Plots Produced Across the Three Packages . . . 136

A.4 Diagnostic Plots Produced With the **BAS** Package 137

List of Tables

2.1	Performance of the BMS, BAS, and BMA Packages for the Fernandez et al. (2001 <i>b</i>) Study.	25
2.2	BMA, WALs, and MMA Estimation for Fernandez et al. (2001 <i>b</i>) Study. . .	27
2.3	BMA, WALs and MMA Estimation for Masanjala & Papageorgiou (2008) Study.	28
2.4	BMA, WALs, MMA Estimation for Doppelhofer & Weeks (2009) Study. . .	29
2.5	BMA, and JMA Estimation for Harrison & Rubinfeld (1978) Study.	35
3.1	Data Descriptive Statistics	47
3.2	Bayesian Model Averaging Analysis: Wave 1	48
3.3	Bayesian Model Averaging Analysis: Wave 2	49
3.4	Bayesian Model Averaging Analysis: Wave 3	50
3.5	Bayesian Model Averaging Analysis: Wave 4	51
3.6	Bayesian Model Averaging Analysis: Waves 1-3	52
3.7	Bayesian Model Averaging Analysis: Waves 1-4	53
3.8	Bayesian Model Averaging Analysis: Wave 1 (No Subjective Variables) . . .	54
3.9	Bayesian Model Averaging Analysis: Wave 2 (No Subjective Variables) . . .	55
3.10	Bayesian Model Averaging Analysis: Wave 3 (No Subjective Variables) . . .	56
3.11	Bayesian Model Averaging Analysis: Wave 4 (No Subjective Variables) . . .	57
3.12	Bayesian Model Averaging Analysis: Wave 1-3 (No Subjective Variables) .	58
3.13	Bayesian Model Averaging Analysis: Wave 1-4 (No Subjective Variables) .	59
3.14	Jointness of Happiness Determinants: Wave 1	60
3.15	Jointness of Happiness Determinants: Wave 1 (Continued)	61
3.16	Jointness of Happiness Determinants: Wave 2	62
3.17	Jointness of Happiness Determinants: Wave 2 (Continued)	63
3.18	Jointness of Happiness Determinants: Wave 3	64
3.19	Jointness of Happiness Determinants: Wave 3 (Continued)	65

3.20	Jointness of Happiness Determinants: Wave 4	66
3.21	Jointness of Happiness Determinants: Wave 4 (Continued)	67
3.22	Jointness of Happiness Determinants: Wave 1-3	68
3.23	Jointness of Happiness Determinants: Wave 1-3 (Continued)	69
3.24	Jointness of Happiness Determinants: Wave 1-4	70
3.25	Jointness of Happiness Determinants: Wave 1-4 (Continued)	71
3.26	Bayesian Model Averaging Analysis: Austria	73
3.27	Bayesian Model Averaging Analysis: Belgium	73
3.28	Bayesian Model Averaging Analysis: Bulgaria	74
3.29	Bayesian Model Averaging Analysis: Switzerland	74
3.30	Bayesian Model Averaging Analysis: Cyprus	75
3.31	Bayesian Model Averaging Analysis: Czech Republic	75
3.32	Bayesian Model Averaging Analysis: Germany	76
3.33	Bayesian Model Averaging Analysis: Denmark	76
3.34	Bayesian Model Averaging Analysis: Spain	77
3.35	Bayesian Model Averaging Analysis: Finland	77
3.36	Bayesian Model Averaging Analysis: France	78
3.37	Bayesian Model Averaging Analysis: Great Britain	78
3.38	Bayesian Model Averaging Analysis: Greece	79
3.39	Bayesian Model Averaging Analysis: Hungary	79
3.40	Bayesian Model Averaging Analysis: Ireland	80
3.41	Bayesian Model Averaging Analysis: Israel	80
3.42	Bayesian Model Averaging Analysis: Italy	81
3.43	Bayesian Model Averaging Analysis: Luxemburg	82
3.44	Bayesian Model Averaging Analysis: Netherlands	82
3.45	Bayesian Model Averaging Analysis: Norway	83
3.46	Bayesian Model Averaging Analysis: Poland	83
3.47	Bayesian Model Averaging Analysis: Portugal	84
3.48	Bayesian Model Averaging Analysis: Russian Federation	84

3.49	Bayesian Model Averaging Analysis: Sweden	85
3.50	Bayesian Model Averaging Analysis: Slovenia	85
3.51	Bayesian Model Averaging Analysis: Slovak Republic	86
3.52	Bayesian Model Averaging Analysis: Turkey	86
3.53	Bayesian Model Averaging Analysis: Ukraine	87
4.1	Variance of Logarithm of Output and Growth Accounting Variables Over Time.	90
4.2	Relative Changes in Variances of Logarithm of Output and Growth Accounting Variables Over Time by Geographical Groups 1960-2000.	98
4.3	Overall, Time and Covariate Effects for Conditional Variance for OECD Countries	99
4.4	Overall, Time and Covariate Effects for Conditional Variance for African Countries	100
4.5	Overall, Time and Covariate Effects for Conditional Variance for Latin American Countries	101
4.6	Overall, Time and Covariate Effects for Conditional Variance for Asian Countries	102
4.7	Overall, Time and Covariate Effects for Conditional Variance for MENA Countries	103
4.8	Conditional Variance Effects by Variable and Region	104
A.1	Grades of Evidence Corresponding to Values of the BIC Difference for M_j Against M_i	125
A.2	A Comparison of BMS, BAS, and BMA Options	127
A.3	A Comparison of Posterior Inclusion Probabilities (PIPs) Across the BMS, BAS, and BMA Packages. (Full Model Enumeration)	129
A.4	A Comparison of Posterior Means and Standard Deviations Across the BMS, BAS, and BMA Packages. (Full Model Enumeration)	130
A.5	Posterior Inclusion Probabilities Across the BMS, BMA, and BAS Packages. (Model Sampling Required)	131
A.6	A Comparison of Posterior Means and Standard Deviations Across the BMS, BMA, and BAS Packages. (Model Sampling Required)	133
A.7	A Comparison of BMS, BAS, and BMA Run Times.	138

A.8	Performance of BMS, BAS, and BMA Packages for the Fernandez et al. (2001 <i>b</i>) Study.	140
A.9	Performance of BMS, BAS, and BMA Packages for the Doppelhofer & Weeks (2009) Study.	141

Chapter 1

Introduction to Model Averaging

1.1 Applied Economics & Model Uncertainty

Applied researchers in economics frequently face the issue of model uncertainty. In general, a lack of consensus regarding theoretical frameworks creates situations where many possible models exist to explain the effect of a particular variable on the outcome of interest. In such a difficult econometric environment, model selection techniques are the most common methods to select the *best* model among a set of candidate models. Once the best model is selected, all subsequent estimations are constructed accordingly, assuming that the selected model is the true data generating process. In empirical analysis, however, picking a single model tends to ignore the uncertainty associated with the specification of the selected model and may result in unstable outcomes and too optimistic confidence intervals. Leeb & Pötscher (2003, 2005) investigate the issues associated with confidence intervals and inference when model selection is applied. An alternative approach to model selection is to estimate all the competing models and compute a weighted average of the estimates. This method, that has become attractive to many practitioners, is called model averaging and is able to incorporate model uncertainty into the analysis.

Empirical growth literature is one the best examples suffering from the issue of model uncertainty. Economic growth—the increase in income per capita—and its causes are conceivably the most substantial areas of research amongst economists. Today, we witness large differences in income per capita and output per worker across nations. Countries at the top of the world income distribution, the United States for example, are nearly 30 times richer than those at the bottom such as Rwanda. Several studies have shown that countries with higher income per capita enjoy a higher quality of life, consumption, standards of living, and life expectancy. Therefore, understanding the economic discrepancies across different nations and understanding the key factors of long-run economic growth are central in macroeconomics. Uncovering the key determinants of economic growth is important because the income per capita gaps have direct welfare consequences and because it will help us understand how different economies function or sometimes fail to function (See Acemoglu 2009). The importance of economic growth and its immense impact on human well-being have allured myriads of economists over the past decades to uncover salient economic growth determinants.

In practice, growth empiricists have been continuously working on deriving statistically powerful regression and distributional-based techniques from theoretical cross-country growth models, including neoclassical and endogenous frameworks, primarily to uncover three pieces of the puzzle: (i) the key elements of long-run economic growth, (ii) parameter heterogeneity

in growth, and (iii) nonlinearities in the growth mechanism. But the open-ended nature of growth theories has created a quite difficult econometric environment for empirical growth researchers (Brock & Durlauf 2001).

By open-endedness, I mean that the validity of one casual theory of growth does not reject the validity of another. For instance, a theory indicating that the rule of law matters for economic growth does not imply the falseness of a theory highlighting a casual relationship between the degree of capitalism and growth. As a result, an individual researcher who is interested in measuring the effect of a particular variable on growth typically considers a single model containing a set of key determinants of growth along with the variable of interest, and then proceeds as if the selected model had generated the data.

As previously mentioned, this approach fails to address the issue of inherent model uncertainty arising in growth regressions. This issue is a fundamental problem facing growth empiricists (Brock & Durlauf 2001), and not taking it into account generates unreliable confidence intervals and statistical inferences for a given phenomenon. Another issue with traditional growth regressions is that many empirical growth studies assume identical parameters, describing growth rate differences across countries. For example, a change in the civil rights index has the same effect on growth rate in the U.K. and Iran according to this assumption—a conclusion that is hard to believe.

In order to properly incorporate model uncertainty and parameter heterogeneity into empirical growth analyses, researchers have used recent theoretical advances and improvements in computational power and speed to implement a variety of model averaging techniques to address both theory uncertainty and parameter heterogeneity, that are the main source of mistrust in cross-country growth regressions. The next section provides a brief introduction to model averaging methods.

1.2 Model Averaging Techniques

Both model uncertainty, that is intrinsic in growth regressions, and the assumption of parameter homogeneity, can be relaxed by treating model specification as an additional unknown feature of the data. This naturally leads to model averaging approaches to data analysis (Durlauf, Kourtellis & Tan 2008). Model averaging involves averaging over all possible combinations of explanatory variables when making statistical inferences about a quantity of interest. Recent studies have shown that model averaging techniques provide better predictive performances than any single selected model in the presence of model uncertainty (Raftery, Madigan & Hoeting 1997). Bayesian model averaging (BMA) and Frequentist model averaging (FMA) are the most promising current approaches to model averaging.

1.2.1 Bayesian Model Averaging

This section briefly reviews the implementation of Bayesian model averaging. Suppose we have k possible explanatory variables, $x_1, x_2, \dots, x_k$ and consequently 2^k possible models, $M = \{M_1, M_2, M_3, \dots, M_{2^k}\}$, that all can potentially explain the data. Considering the j th model, M_j , a generic linear regression model under M_j is

$$y = X^j \beta^j + \epsilon \quad (1.1)$$

where y is the vector of observations on the variable of interest, X^j is the matrix of all explanatory variables considered for the j th model, β^j is the vector of parameters of model M_j , and ϵ is the standard error term. The probability distribution of β^j given j th model, M_j and data, Δ , is

$$p(\beta^j | \Delta, M_j) = p(\Delta | \beta^j, M_j) \frac{p(\beta^j | M_j)}{p(\Delta | M_j)} \quad (1.2)$$

This probability distribution can be used to measure the probability that M_j actually generated the data. Using Bayes' rule

$$p(M_j | \Delta) = p(\Delta | M_j) \frac{p(M_j)}{p(\Delta)} \quad (1.3)$$

where $p(M_j)$ is the prior model probability, and $p(\Delta | M_j)$ is the marginal or integrated likelihood. Practical use of this equation requires few manipulations. As is apparent, $p(M_j)$ does not depend on the data and essentially shows how likely we believe M_j to be the correct model before seeing the data. To calculate the marginal likelihood, one can integrate both sides of equation (1.2) with respect to β^j that yielding

$$\begin{aligned} \int p(\beta^j | \Delta, M_j) d\beta^j &= \int p(\Delta | \beta^j, M_j) \frac{p(\beta^j | M_j)}{p(\Delta | M_j)} d\beta^j \\ \Rightarrow 1 &= \frac{1}{p(\Delta | M_j)} \int p(\Delta | \beta^j, M_j) p(\beta^j | M_j) d\beta^j \\ \Rightarrow p(\Delta | M_j) &= \int p(\Delta | \beta^j, M_j) p(\beta^j | M_j) d\beta^j \end{aligned} \quad (1.4)$$

where $p(\beta^j | M_j)$ is prior probability distribution assigned to the parameters of model M_j . Whether or not the data favors M_i over M_j depends on the Bayes factor, $\frac{p(\Delta | M_i)}{p(\Delta | M_j)}$. When this ratio is greater than 1, then the data supports M_i over M_j . Finally, $p(\Delta)$, the last piece of equation (1.3), can be computed using the law of total probability as follows

$$p(\Delta) = \sum_{i=1}^{2^k} p(\Delta | M_i) p(M_i) \quad (1.5)$$

The logic of BMA, then, suggests that the posterior density of $\beta = (\beta_1, \beta_2, \dots, \beta_k)'$ in full model $y = X\beta + \epsilon = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \epsilon$ is

$$p(\beta|\Delta) = \sum_{i=1}^{2^k} p(\beta|\Delta, M_i) p(M_i|\Delta) \quad (1.6)$$

Taking expectations across (1.6) gives the posterior mean for parameter β as

$$E(\beta|\Delta) = \sum_{i=1}^{2^k} E(\beta|\Delta, M_i) p(M_i|\Delta) \quad (1.7)$$

Following Leamer (1978), the posterior variance of parameter of interest would be

$$Var(\beta|\Delta) = \sum_{i=1}^{2^k} Var(\beta|\Delta, M_i) p(M_i|\Delta) + \sum_{i=1}^{2^k} (E(\beta|\Delta, M_i) - E(\beta|\Delta))^2 p(M_i|\Delta) \quad (1.8)$$

A useful feature of BMA is that it allows the estimation of the posterior probability that a particular explanatory variable, let say x_h , belongs to the true model. This probability is the sum of the posterior probabilities for all the models that include x_h . Mathematically,

$$p(x_h|\Delta) = \sum_{i: x_h \in M_i} p(M_i|\Delta) \quad (1.9)$$

If posterior inclusion probability (hereafter PIP) obtained from (1.9) is high for variable x_h , this variable is considered as a robust determinant of variable of interest.

While BMA is an intuitively attractive solution to be used to overcome the problem of model uncertainty in empirical research, its implementation presents a number of difficulties and challenges. For instance, the size of model space increases exponentially as the number of candidate predictors increases, making the summation in equation (1.6) computationally infeasible. Another difficulty is that the integrals used for the marginal likelihood of models can be hard to compute unless the model classes are linear regressions (Raftery et al. 1997) or discrete graphical (Madigan, York & Allard 1995) where closed form solutions are available.

To overcome the problem of exploding model space size in the presence of large numbers of regressors, two approaches are common. The first method involves applying the Occam's window algorithm presented in Madigan & Raftery (1994) to linear regression models. This algorithm discards all the models that predict the data far less accurately than the models which provide the best predictions. This means all models belonging to

$$\mathcal{A} = \left\{ M_j : \frac{\max_l \{p(M_l|D)\}}{p(M_j|D)} > C \right\} \quad (1.10)$$

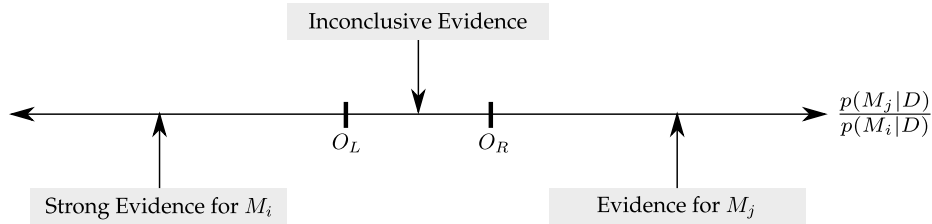
will be excluded where C is the data analyst's choice and $\max_l \{p(M_l|D)\}$ represents the model with the highest posterior model probability. This algorithm also excludes all com-

plex models that receive less support from the data than their simpler counterparts. Mathematically, models that are in

$$\mathcal{B} = \{M_j : \exists M_l \in \mathcal{A}', M_l \subset M_j, p(M_j|D) < p(M_l|D)\} \quad (1.11)$$

will not be considered. These two principles remarkably reduce the model space size. Now all we need is a search mechanism to identify the models in $\mathcal{A}$. Having said that, Occam's window search algorithm looks at the ratio of posterior model probabilities $p(M_j|D)/p(M_i|D)$. Here M_j has one less predictor than M_i . Figure 1.1 presents the idea.

Figure 1.1: A Visualization of Occam's Window



If there is evidence for model M_j , then this algorithm rejects M_i , but to reject model M_j , which is less complex than M_i , we need strong evidence for the larger model, M_i . And if the evidence is inconclusive, we have fallen in Occam's window and neither model is rejected. Madigan & Raftery (1994) initially chose 1/20 and 1 for O_L and O_R , respectively but Raftery, Madigan & Volinsky (1996) showed that 1/20 and 20 as the choice for O_L and O_R , respectively, can improve the predictive performance. The shortcoming of Occam's window method is that it accounts for model uncertainty by averaging only over a restricted class of models that may lead to biased posterior probabilities. It is especially the case when there are many similar models that dilute a substantial posterior probability, but the algorithm tends to reject them.

An alternative to Occam's window approach is to use Markov Chain Monte Carlo sampler (Tanner & Wong 1987, Gelfand 1990, Smith & Roberts 1993, Besag, Green, Higdon & Mengersen 1995, Madigan et al. 1995) that is the most common method used to obtain samples from the posterior distributions. Markov Chain Monte Carlo (MCMC) methods sample from probability distributions based on constructing a Markov chain that has the desired distribution as its equilibrium distribution. The state of the chain after a large number of steps is then used as a sample of the desired distribution. The quality of the sample improves as the number of steps increases.

MCMC methods typically move around the equilibrium distribution in relatively small steps and can take a long time for the sampler to explore all of the space. To walk through model space using MCMC methods, different walkers are proposed. Metropolis-Hastings algorithm (Metropolis, Rosenbluth, Rosenbluth, Teller & Teller 1953, Hastings 1970), Gibbs sampling (Geman & Geman 1984, Casella & George 1992), Slice sampling (Neal 2003), and Multiple-try Metropolis (Liu, Liang & Wong 2000) are some examples. Researchers have also proposed more sophisticated algorithms that exhibit faster convergence by preventing the sampler from doubling back to the model space, but these methods, compared

to their normal peers, are harder to implement. If the models have parameter spaces of different dimensions, the reversible-jump method proposed in Green (1995) that is a variant of Metropolis-Hastings algorithm is applicable.

Other controversial aspects of implementing BMA include the specification of the prior model probabilities $p(M_j)$ as well as the parameters' prior probability distributions, $p(\beta^j|M_j)$. The set of prior probabilities must be informative with respect to the likelihood, meaning priors should be relatively high where the likelihoods are large; otherwise, the choice of priors will have a substantial effect on the posteriors.

Several remedies exist in the literature for the specification of model priors $p(M_j)$. The most common model prior structure in BMA analysis is the Binomial distribution. According to this structure, the model size (ϑ) follows a Binomial distribution with probability of success θ :

$$\vartheta \sim \text{Bin}(\ell, \theta) \quad (1.12)$$

where ℓ is the number of regressors included in the model and θ is the prior inclusion probability assigned to each variable. Considering Binomial distribution structure, the prior probability attached to the j th model with ℓ_j regressors is given by

$$p(M_j) = \theta^{\ell_j} (1 - \theta)^{\ell - \ell_j} \quad (1.13)$$

A special case of this prior structure is to assume $\theta = 1/2$, implying that the researcher has no preference for a specific model. That is, the prior probability assigned to each model will be

$$p(M_1) = p(M_2) = \dots = p(M_{2^\ell}) = 1/2^\ell \quad (1.14)$$

It is worth mentioning that Fernandez, Ley & Steel (2001a) used this uniform distribution over the model space to investigate the main determinants of economic growth. An alternative prior specification, proposed by Ley & Steel (2009a), is Binomial-Beta structure in which θ is also treated as a random value. This additional randomness implies an increase in prior uncertainty about the model size (ϑ) and makes the choice of model priors less critical. In particular, the model size and the prior inclusion probability for each variable are distributed as the following

$$\vartheta \sim \text{Bin}(\ell, \theta) \quad (1.15)$$

$$\theta \sim \text{Bet}(\lambda, \mu) \quad (1.16)$$

where λ and μ are hyper-parameters that will be set by the researcher. Combining equations (1.15) and (1.16), Ley & Steel (2009a) show that

$$E(\vartheta) = \frac{\lambda}{\lambda + \mu} \ell \quad (1.17)$$

and propose to set $\lambda = 1$ and $\mu = (\ell - E(\vartheta))/E(\vartheta)$. This means we only need to specify the expected prior model size in this setup. A shortcoming of both the Binomial and Binomial-Beta priors is that they assume the probability of one explanatory variable appearing in the model is independent of other variables, which is usually violated in empirical work

because regressors are typically correlated. In other words, the Binomial & Binomial-Beta distributions would not penalize the inclusion of redundant variables in the model. This issue was first highlighted by George (1999) and called dilution problem. Durlauf et al. (2008) solved the dilution problem by assigning probabilities to neighborhoods of models rather than models themselves. Their method is designed to penalize models with many redundant variables.

Prior density choice for BMA analysis is not limited to model priors. Priors on the parameter space also need to be specified. Most of BMA studies assume a conditionally normal distribution as the choice of prior structures for the coefficients with zero mean and a variance structure proposed by Zellner (1986). More specifically, I assume

$$p(\beta^j | M_j) \sim N(0, \sigma^2 \Omega) \quad (1.18)$$

where prior variance matrix Ω is assumed to be proportional to the sample covariance,

$$\Omega = (gX^{j'}X^j)^{-1} \quad (1.19)$$

The form of the factor of proportionality in this setup g , which was first suggested by Zellner (1986), is crucial in BMA analyses. For example, Fernandez et al. (2001a) recommend setting $g = n$, or $g = \max(n, \ell^2)$ where n is the sample size and ℓ is the number of regressors, Foster & George (1994) introduce a risk inflation criterion and set $g = \ell^2$, and George & Foster (2000) and Liang, Paulo, Molina, Clyde & Berger (2008) estimate a local empirical Bayes g -parameter. Eicher, Papageorgiou & Raftery (2011) conduct a thorough study of different prior structures and show that the combination of the Unit Information Prior (UIP), initially proposed by Kass & Wasserman (1995), on the parameter space and the uniform distribution on the model space is superior to any other possible combinations of priors proposed in the BMA literature. Throughout this work, I follow their suggestion for the choice of priors on parameters and models to conduct empirical BMA analysis. Chapter 3 of this work provides an empirical application of BMA framework to capture the main determinants of life satisfaction across the European Union. In addition, Appendix A outlines in detail the implementation of a Bayesian model averaging exercise, using the computational power of computers coupled with the power of MCMC machinery in computing multidimensional complex integrals.

Although BMA has become a mainstream tool in empirical settings with large numbers of potential regressors, it remains problematic. The set of priors on parameters within a model and the set of the prior model probabilities must be specified before calculating posterior probabilities attached to different models. As stated in Brock, Durlauf & West (2003) and Ley & Steel (2009a), the choice of prior distributions can be extremely critical for the outcome of BMA analysis. These prior probabilities must be informative with respect to the likelihood, meaning priors should be relatively high where the likelihoods are large; otherwise, the choice of priors will have a substantial effect on the posteriors. Another issue, raised by Hjort & Claeskens (2003), is the difficulty of dealing with the priors when they are in conflict with each other, stemming from mixing together many prior opinions regarding the parameters of interest. Non-Bayesian approaches, usually called Frequentist model averaging (FMA), are other solutions to model uncertainty. From a Frequentist perspec-

tive, uncovering information about a particular parameter of interest involves calculating an estimate of the parameter of interest along with an associated standard error for the estimate through assigning appropriate weights to each model. Section 1.2.2 summarizes the implementation of FMA.

1.2.2 Frequentist Model Averaging

Model averaging from a Frequentist perspective had not received widespread attention until recent theoretical advances that have facilitated their practical implementation. In contrast with the BMA approach, the FMA approach requires no priors on parameters and models and the corresponding estimators are determined exclusively by data. Unlike the BMA approach, FMA is less complex and can be implemented without much difficulty or dramatically prolonged and intensive computations. The formal definition of FMA estimator for a linear model is as follows (The discussion can be extended to more general models). Assume a linear regression model, given data D

$$y = X\beta + \epsilon = Z\gamma + W\theta + \epsilon \quad (1.20)$$

where $y(n \times 1)$ is the vector of dependent variable, $X = [Z : W]$ is the matrix of explanatory variables, $\beta = (\gamma, \theta)$ is the parameter vector, and $\epsilon(n \times 1)$ is the standard random error term. Here, Z contains the explanatory variables that are required to be in the model on theoretical grounds, while W contains additional explanatory variables about which there is doubt as to whether or not they should be in the model. The FMA estimator of β is then given by

$$\hat{\beta} = \sum_{j: M_j \in M} \omega_j \hat{\beta}^j \quad (1.21)$$

where $M = \{M_1, M_2, \dots, M_K\}$ denotes the model space, $\hat{\beta}^j$ is the estimation of parameter of interest under model M_j , and ω_j is the weight assigned to the model, M_j . The weights are determined based on the data and no priors are involved. For asymptotic optimality purposes, I require the weight vector, $\omega = (\omega_1, \omega_2, \dots, \omega_j, \dots, \omega_K)$, to be in the unit simplex in $\mathbb{R}^K$:

$$\mathcal{U} = \left\{ j : \omega_j \in [0, 1], \sum_{j=1}^K \omega_j = 1 \right\} \quad (1.22)$$

A proper way of computing the variance of FMA estimators is still an open area of research as a formal statistical framework on FMA inference is not fully developed yet. Buckland, Burnham & Augustin's (1997) approach is perhaps the most widely-used method when constructing confidence intervals of FMA estimators. They propose the following formula

$$\text{Var}(\hat{\beta}) = \sum_{j=1}^K \omega_j \sqrt{\text{Var}(\hat{\beta}^j | \text{bias}_j) + \text{bias}_j^2} \quad (1.23)$$

that accounts for the additional uncertainty due to model selection. In this formula $\text{bias}_j = \hat{\beta}^j - \hat{\beta}$ is the misspecification bias that arises in estimating β under model j . In practical work, the vector of parameters, β^j , under model j , and also the variances of estimated parameters is calculated then the vector of estimated parameters under model j will be subtracted from the estimates obtained from the full model. In this setup, the full model, including all covariates, is used as the benchmark to calculate the misspecification bias.

The last piece of information I need in order to implement the FMA estimator is to construct suitable weights that are supported by statistical theories. In recent years, several ad hoc weight choice methods have been proposed within Frequentist framework that are asymptotically optimal and achieve the lowest possible squared error among the same class of estimators. For example, Buckland et al. (1997) construct model weights based on different information criteria values obtained from the competing models¹. They propose using the following weights

$$\omega_j = \frac{\exp(-I_j/2)}{\sum_{i=1}^K \exp(-I_i/2)} \quad (1.24)$$

where

$$I_j = -2\log(\Gamma_j) + \Xi_j \quad (1.25)$$

and Γ_j is the maximum likelihood function for the j th model and Ξ_j is a penalty term of the number of regressors used in the j th model. Two common choices for the penalty function are $\Xi_j = 2\ell_j$ and $\Xi_j = \ell_j \ln(n)$ where ℓ_j is the number of explanatory variables in model j and n is the sample size. The former choice corresponds to the Akaike Information Criterion (AIC), introduced by Akaike (1973), and the latter one corresponds to the Bayesian Information Criterion (BIC).

Yang (2001) employed an adaptive regression using a mixing method. This weighted estimator combines the estimators from different regression procedures including simple linear regression, additive modeling, projection pursuit, neural network, and so on. Yuan & Yang (2005) used a model screening step in conjunction with this method. Lately, Leung & Barron (2006) proposed a mixture least squares estimator withing Frequentist framework using a vector of weights that depends on the estimator's risk characteristics, while Claeskens & Hjort (2008) proposed selecting model weights based on the Focused Information Criterion (FIC) originally presented in Claeskens & Hjort (2003).

Recently, Hansen (2007) proposed a least squares model averaging estimator with model weights selected by minimizing the Mallows' criterion described in Mallows (1973). This estimator, known as Mallows model averaging (MMA), is easily implementable for linear regression models and has very desirable asymptotic optimality properties. Its optimality lies in the fact that the Mallows' criterion is asymptotically equivalent to the squared error. Therefore, the MMA estimator that minimizes the Mallows' criterion also minimizes the squared error in large samples. Section 2.3.3 of chapter 2 shows how this estimator can be applied to growth regressions to assess the key determinants of economic growth. The results are compared to those obtained from Bayesian model averaging paradigm.

¹This idea was extended further in Burnham & Anderson (2002, Chapter 6)

To implement the MMA approach laid out in Hansen (2007) consider a standard linear regression model

$$y = X\beta = \sum_{i=1}^{\infty} \beta_i x_i + \varepsilon \quad (1.26)$$

where y is a $n \times 1$ vector of dependent observations, ε is a homoskedastic error term assumed to be distributed independently with mean 0 and variance σ^2 . Let X_{k_j} be the corresponding submatrix of X containing the first k_j columns of X .² Index $j = 1, 2, 3, \dots, K$ indicates the sequence of models. I estimate the ℓ th model, $y = X_{k_\ell} \beta^{(\ell)} + b_{k_\ell} + \varepsilon$, by least squares and denote the vector of estimated coefficients under model ℓ as $\hat{\beta}^{(\ell)}$ and the vector of estimated residuals as $\hat{\varepsilon}_\ell$, such that $b_{k_\ell} = \sum_{j=k_\ell+1}^{\infty} \beta_j x_{ij}$ is the approximation error of the ℓ th model.

This procedure is repeated for all K models and the final estimator of β is

$$\hat{\beta} = \sum_{j=1}^K \omega_j \begin{pmatrix} \hat{\beta}^{(j)} \\ 0 \end{pmatrix} \quad (1.27)$$

where ω_j , $j = 1, 2, \dots, K$ are normalized weights calculated numerically by minimizing the Mallows criterion. This criterion can be presented as

$$C_p(\Omega) = \Omega' \hat{\varepsilon}' \hat{\varepsilon} \Omega + 2\sigma^2 L' \Omega \quad (1.28)$$

where $\Omega_{(K \times 1)} = (\omega_1, \omega_2, \dots, \omega_K)'$ is the vector of weights, $\hat{\varepsilon}_{(n \times K)} = (\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_K)$ is the matrix of all residual vectors, and $L_{(K \times 1)} = (k_1, k_2, \dots, k_K)'$ is the vector of the number of regressors used for each of the K models.³ This is, in fact, a classic linear-quadratic problem. Mathematically, the optimization problem is

$$\begin{aligned} \min_{\omega_j} C_p(\Omega) &= \Omega' \hat{\varepsilon}' \hat{\varepsilon} \Omega + 2\sigma^2 L' \Omega \\ \text{s.t.} \quad \sum_{j=1}^K \omega_j &= 1, \quad 0 \leq \omega_j \leq 1, \quad \forall j = 1, 2, \dots, K. \end{aligned} \quad (1.29)$$

MMA is asymptotically an optimal estimator if the error terms are homogeneous. In heteroskedastic error settings, Hansen & Racine's (2009) Jackknife model averaging estimator (JMA) is a desirable choice. They propose selecting the model weights by minimizing a delete-one cross-validation criterion. JMA is the first Frequentist estimator that allows for heteroskedasticity and achieves the lowest possible expected squared error over the class of linear estimators. In nutshell, suppose that a sample of n observations $(y_{i(1 \times 1)}, x_{i(1 \times k)}), i = 1, 2, \dots, n$, is given and one believes that data generating process is $y_i = \mu_i + e_i$, where μ_i is the conditional mean of y_i given x_i , and e_i is a heteroskedastic error term such that $E(e_i | x_i) = 0$ and $E(e_i^2 | x_i) = \sigma_i^2$. Let define the vector of observations on independent variable as $y_{n \times 1} = (y_1, y_2, \dots, y_n)'$, the vector of fitted values as $\mu_{n \times 1} = (\mu_1, \mu_2, \dots, \mu_n)'$, the vector of error terms as $e_{n \times 1} = (e_1, e_2, \dots, e_n)'$,

²Since I assume $E(\varepsilon | X) = 0$ therefore the first column of X is a vector of ones representing the intercept.

³A thorough review of MMA is provided in Wang, Zhang & Zou (2009).

and the matrix of regressors as $X_{n \times k} = (x_1, x_2, \dots, x_n)'$. If there are M model candidates $\{\mu^1, \mu^2, \dots, \mu^M\}$ to estimate the fitted values, the JMA estimator of μ takes the form $\mu = \sum_{j=1}^M \omega_j \mu^j$, where $\omega_{(1 \times 1)}^j$ is the weight assigned to the j th model using JMA approach and $\mu^j = (\mu_1^j, \mu_2^j, \dots, \mu_n^j)'$ is the vector of fitted values obtained from the j th model such that μ_i^j is computed with the i th observation deleted (known as leave-one-out cross-validation). For instance, if one uses OLS to estimate each model, then

$$\mu_i^j = x_i^{j'} (X_{(-i)}^{j'} X_{(-i)}^j)^{-1} X_{(-i)}^{j'} y_{(-i)} \quad (1.30)$$

where $X_{(-i)}^j$ and $y_{(-i)}$ represents matrices $X_{n \times k}^j$ and $y_{n \times 1}$ with the i th row deleted. Consequently, the jackknife residual vector of the j th model is $e_{(n \times 1)}^j = y_{(n \times 1)} - \mu_{(n \times 1)}^j$ and the $n \times M$ matrix of residuals is $\Upsilon_{n \times M} = (e_{n \times 1}^1, e_{n \times 1}^2, \dots, e_{n \times 1}^M)$. The JMA choice of optimal weights minimizes

$$\begin{aligned} \min_{\omega_j} CV(\Omega) &= \frac{1}{n} \Omega' \Upsilon' \Upsilon \Omega \\ \text{s.t.} \quad \sum_{j=1}^M \omega_j &= 1, \quad 0 \leq \omega_j \leq 1, \quad \forall j = 1, 2, \dots, M. \end{aligned} \quad (1.31)$$

where $\Omega_{(M \times 1)} = (\omega^1, \omega^2, \dots, \omega^M)'$ is the weight vector consisting of the weights assigned to the M models. Section 2.6 of chapter 2 reviews in details the implementation of this estimator, provides an empirical example, and proposes an econometric technique to overcome the issue of having a large model space.

In closing, it is worth mentioning that FMA has been recently used in many empirical studies in different fields of science. For instance, Claeskens, Croux & Kerckhoven (2006) applied a Frequentist model averaging approach to a study of diabetic retinopathy (Disease of the retina that results in impairment or loss of vision). Pesaran, Schleicher & Zaffaroni (2009) employ FMA approach in risk management and provide an application to futures markets. And Wan, Zhang & Zou (2010) use FMA as a statistical tool to evaluate the economic impacts of tourism industry. The interested readers are referred to Hjort & Claeskens (2003) and Wang et al. (2009) for a comprehensive review of this literature.

As stated earlier in section 1.1, theory uncertainty and heterogeneity uncertainty are not the only source of mistrust in empirical growth regressions. Nonlinearity and functional form uncertainty are other issues that cast doubt on growth empirics. Model averaging techniques fully address the theory uncertainty issue and partially address parameter heterogeneity but these techniques typically assume a log-linear growth regression so that functional form uncertainty can be repealed. Cobb-Douglas or Constant Elasticity of Substitution (CES) functions are common functional forms used in the study of economic growth. However, there have been some recent studies that reject the assumption of linearity and prove the presence of nonlinearities in cross-country regressions allowing for multiple regimes of growth patterns among different countries. Maasoumi, Racine & Stengos (2007) and Fiaschi & Lavezzi (2003) are recent studies showing that growth appears as a nonlin-

ear process. Failure to account for nonlinearity in empirical growth regressions may lead in misleading conclusions. For example, the presence of nonlinearities would cast doubt on the studies aiming to estimate the speed on convergence based on a concave production function. To capture nonlinearities and parameter heterogeneity in empirical growth practice, I use recently-developed methods for nonparametric regression. The next section provides a brief review of nonparametric methods and their application in growth empirics.

1.3 Nonparametric Statistics for Growth Empirics

The prominence of nonparametric or distribution-free methods in modern statistics and econometrics has grown rapidly since their inception in the mid-1930s. Requiring few or no stringent assumptions about the populations from which data are obtained, these highly efficient techniques have emerged as the preferred methodology among researchers performing data analysis. Nonparametric methods are robust to functional form misspecification and have the ability to remove irrelevant variables from the model. That being said, suppose we have n independent and identically distributed observations $(y_i, x_{i1}, x_{i2}, \dots, x_{ik})$, $i = 1, 2, \dots, n$ from an unknown continuous distribution with density $f(y, x_1, x_2, \dots, x_k) = f(y, x)$. Now consider a growth regression model

$$y_i = m(x_i) + u_i, \quad i = 1, 2, \dots, n \quad (1.32)$$

where y_i is the dependent variable, here growth rate, x_i is the vector of k potential regressors, u_i is the disturbance terms, and $m(x_i) = E(y|x_i)$ is an unspecified conditional mean of y_i . Then in a parametric universe, we make a good many assumptions about the model. We assume the form of $m(x_i)$ is known, usually linear in x_i , i.e. $m(x_i) = E(y|x_i) = \alpha + \beta'x_i$ or log-linear in x_i , i.e. $m(x_i) = E(y|x_i) = \alpha + \beta' \log x_i$, and we assume the conditional variance of y_i given x_i and the joint density of y_i and x_i are known.

In a nonparametric universe, on the other hand, my goal is to estimate the unknown regression function $m(\cdot)$ without making explicit assumptions about its functional form. To do so, I use the insights of Nadaraya (1964), Watson (1964), Stone (1977), Cleveland (1979), and Hall, Li & Racine (2007) to construct my nonparametric estimator. Suppose that the elements of x_i vector are divided to continuous regressors, x_i^c , unordered discrete regressors, x_i^u , and ordered discrete regressors, x_i^o . The local constant least squares (LCLS) estimator of $m(\cdot)$ begins by minimizing

$$\min_m \sum_{i=1}^n (y_i - m) K_h(x, x_i) \quad (1.33)$$

This minimization yields

$$\hat{m}(x) = \sum_{i=1}^n K_h(x, x_i) y_i \bigg/ \sum_{i=1}^n K_h(x, x_i) \quad (1.34)$$

where

$$K_h(x, x_i) = \prod_{s=1}^p h_s^{-1} k^c \left(\frac{x_s^c - x_{si}^c}{h_s} \right) \prod_{s=1}^q k^u(x_s^u, x_{si}^u, \lambda_s^u) \prod_{s=1}^r k^o(x_s^o, x_{si}^o, \lambda_s^o) \quad (1.35)$$

Here, $K_h(x, x_i)$ is the commonly used product kernel function, k^c is the standard normal kernel function, k^u is the Aitchison & Aitken's (1976) kernel function, k^o is the Wang & Van Ryzin's (1981) kernel function, and h_s , λ_s^u , λ_s^o are the window widths that determine how many of x_i s around x are to be used in constructing the kernels used to weight observations. For the values of x_i that lie far from x , the assigned weight will be small. Thus, these observations play less important role in the determination of $m(\cdot)$. Having said that, the shape of the kernel functions is not very important factor in estimating the fitted values. Gaussian, Epanechnikov and Uniform kernels are standard functions widely used in the literature. A nice property of the LCLS estimator is its ability to detect irrelevant regressors by assigning large bandwidths to those variables. Another nonparametric estimator of $m(\cdot)$, known as local linear least squares (LLLS), begins by minimizing

$$\min_{\{m, \beta\}} \sum_{i=1}^n (y_i - m - (x_i - x)\beta)^2 K_h(x_i, x) \quad (1.36)$$

Expanding $m(x_i)$ around x by taking a first-order Taylor expansion yields

$$m(x_i) \approx m + (x_i - x) \frac{\partial m}{\partial x}(x^c) = m + (x_i^c - x^c) \beta(x^c) \quad (1.37)$$

Substituting the Taylor expansion back to the equation 1.36 and minimizing it with respect to m and β gives the LLLS estimator of $\delta(x) \equiv (m(x), \beta(x^c))'$ which can be expressed as

$$\hat{\delta}(x) = \left[\sum_{i=1}^n K_h(x, x_i) \begin{pmatrix} 1 \\ x_i^c - x^c \end{pmatrix} (1, (x_i^c - x^c)') \right]^{-1} \sum_{i=1}^n K_h(x, x_i) \begin{pmatrix} 1 \\ x_i^c - x^c \end{pmatrix} y_i \quad (1.38)$$

A nice property of LLLS estimator is its ability to detect whether a specific variables enters the model in a linear fashion or in a nonlinear fashion when automated bandwidth selection is used. The key aspect of any kernel-based estimation is the choice of bandwidths. Too small bandwidths lead in undersmoothed estimations with low bias and high variance while too large bandwidths will lead in having oversmoothed estimations with high bias and low variance. Hence, using methods that select value for bandwidths by automatic means are more desirable. All automatic selection procedure inevitably involves optimizing a criterion. Although many proposals in the literature, Least Squares Cross-Validation (LSCV) procedure, suggested by Rudemo (1982) and Bowman (1984), is my choice of bandwidth selection procedure. The LSCV (leave-one-out) bandwidths are obtained by minimizing the following mean integrated squared error (MISE) term with respect to h

$$MISE = E \left[\int_{-\infty}^{\infty} (\hat{m}(x) - m(x))^2 dx \right] \quad (1.39)$$

Hall et al. (2007) show that the LSCV technique, in a local constant least squares framework, has the ability to smooth out irrelevant variables that have been included in the model and in the local linear least squares framework has the ability to detect the variables that enter in model linearly. More discussions on the choice of window widths can be found in Pagan & Ullah (1999) and Li & Racine (2007).

Chapter 4 of this study provides an application of nonparametric methods in growth empirics. My goal is to find the reasons behind the fact that the variation in output per capita across countries has been constantly increasing over the past decades. I set up my model in a conditional setting using robust, nonparametric estimators and propose a decomposition of the conditional variation of per capita output. As mentioned earlier, the advantage of using nonparametric framework is that I do not have to make assumptions on either the functional form of global production functions or the distribution of the error terms. In this chapter, I take a distributional approach to empirical growth and focus on the shape of the world income distribution as a whole. Assuming that contemporary income differences between countries are transitory, I am looking to find if the dispersion of these differences will decline across time. One limitation of the distributional approach to convergence is that it is not clear how one can formulate a sensible notion of conditional convergence when the second moment of distribution is the parameter of interest (Durlauf, Johnson & Temple 2005). I believe that using nonparametric methods is the right path toward systematically formulating the change in variance of the conditional distribution of output per capita.

Chapter 2

Comparisons of Frequentist Model Averaging Techniques

2.1 Abstract

The primary aim of this paper is to shed more light on the issue of parameter estimation under model uncertainty. Model uncertainty consists of three main types: theory uncertainty, which focuses on which principal determinants should be included in a model; heterogeneity uncertainty, which relates to whether or not the parameters that describe dependent variable are identical across observations; and functional form uncertainty or nonlinearity, which relates to which regressors enter in the model linearly and which ones enter in a non-linear fashion. While Bayesian approaches are dominant methods used to average over the model space, I take a fresh look into Frequentist model averaging techniques and propose statistical routines that computationally ease the implementation of these methods. I look at three important studies on growth theory uncertainty that employed Bayesian model averaging tools. I compare these results with estimates obtained using recently developed Frequentist and alternative Bayesian model averaging techniques. Overall, I successfully replicate all three studies using freely available software in the statistical environment R, provide an easily implementable algorithm to operationalize the Frequentist model averaging methods and find that the sign and magnitude of these new estimates are reasonably close to those produced via traditional Bayesian methods.

2.2 Introduction

The intrigue of uncovering a singular process that dictates economic growth has driven myriad economists to postulate growth theories explicating a link between a set of ‘determinants’ and a measure of economic growth. The propagation of theories on economic growth has simultaneously put forth an empirical conundrum known as “theory openendedness” (Brock & Durlauf 2001). Theory openendedness suggests that while numerous theories may indeed explain cross country growth, no one theory evinces a clear link with economic growth in the face of alternative growth determinants (both within and across growth theories). The lack of empirical robustness of growth models has sharpened the need for viable model selection and averaging techniques to parse through the vast amounts of data being used to empirically test competing growth models. Such techniques allow growth empiricists to tenably focus on the (set of) variables which produce robust relationships with economic growth.

For example, on a review of the literature considering alternative growth theories Durlauf et al. (2005) list over 140 different determinants which have been used in attempts to explain economic growth. The sheer volume of alternative regressors coupled with the limited number of countries presents quite a difficult econometric environment for growth empiricists.¹ To combat the array of potential growth determinants a common approach has been to employ model averaging econometric tools. Specifically, growth empiricists have used advances in both computational power and speed to implement a variety of Bayesian model averaging (BMA) tools to uncover the effect that a given variable has on economic growth along with the posterior probability that a specific variable resides in the final ‘model’ explaining growth. Empirically, BMA methods have proved quite successful at delineating robust growth determinants and theories (see Durlauf et al. 2008).

However, for those more accustomed to operating within the Frequentist paradigm, the appeal of BMA tools is diminished. Until recently, there existed a dearth of available Frequentist model averaging (FMA) tools to directly compete with BMA. Hansen (2007) developed an easily implementable model averaging tool for linear regression known as Mallows model averaging (MMA) that has very desirable asymptotic optimality properties. These methods encapsulate model averaging within the Frequentist econometric arena and are comparable to BMA tools. Thus, I seek to employ MMA to see how my results differ from similar analyses conducted using BMA within the growth theory uncertainty setting. Furthermore, I also consider a modified Bayesian estimator developed by Magnus, Powell & Prüfer (2010) which uses Bayesian insights to average over a set of Frequentist estimators. This estimator is termed weighted average least squares (WALS).

My focus will have several specific aims. First, I replicate the key results of three previous studies using BMA to evaluate various aspects of growth theory uncertainty: Fernandez et al. (2001*b*), Masanjala & Papageorgiou (2008) and Doppelhofer & Weeks (2009). This replication is important given the availability of different software to conduct BMA. Specifically, I focus on replicability within the computing environment R (R Development Core

¹Not to mention additional econometric issues such as endogeneity, functional form specification, measurement error, etc.

Team 2010). Within R there exist at least three alternative packages² to conduct BMA, the **BMS** (Feldkircher & Zeugner 2009), **BMA** (Raftery, Hoeting, Volinsky, Painter & Yeung 2010), and **BAS** (Clyde 2010) packages. Replicability of empirical studies using BMA tools is important since the search algorithms deployed can be computationally involved and it is important to ensure that the final results presented are inherent of the underlying relationship and not that of the search mechanism.

I find that model averaging using the **BMS** package exactly replicates the results of Fernandez et al. (2001b) and Doppelhofer & Weeks (2009) as well as the working paper version of Masanjala & Papageorgiou (2008). The **BMA** package is able to exactly replicate the results of Masanjala & Papageorgiou (2008) but neither of the other two studies, while the **BAS** package cannot replicate the results of any of the three studies.

Second, I directly compare the BMA posterior means and standard deviations of my three studies of interest with the model averaging methods (MMA and WALS) described above. I find that my final weighted estimated parameter vectors using both MMA and WALS are nearly identical to the posterior means obtained using BMA. However, for each of the three examples I encounter differences between the new methods' standard errors and the standard deviation of the posterior distribution of each parameter stemming from the original BMA analysis. The difference in standard errors in the MMA setting most likely owns to the fact that a formal asymptotic theory does not yet exist for the MMA estimator and the method to construct standard errors are conservative. Third, I show that the model enumeration property of the WALS estimator (linear in the number of covariates as opposed to exponential) also translates to the MMA estimator. Given that no appropriate sampling mechanism from the model space exists for the MMA estimator, this enumeration property makes implementation of the MMA estimator trivial for the number of covariates typically encountered in econometric exercises.

The remainder of my paper is as follows. Section 2.3 discusses the implementation of model averaging estimators across the menu of available methods. Sections 2.4 and 2.9 briefly describe the data used for my replications. Section 2.5 and 2.10 present my results for the BMA, WALS, MMA, and JMA estimators and discuss differences and similarities across the estimates while Section 2.11 concludes.

2.3 Model Averaging Techniques

2.3.1 Bayesian Model Averaging

Given the prevalence of descriptive overviews on BMA (Raftery 1995, Hoeting, Madigan, Raftery & Volinsky 1999), I only briefly describe its implementation here. A BMA analysis begins by considering a set of possible models, say $M = \{M_1, M_2, \dots, M_K\}$, where M represents the model space the averaging will be conducted over. Once this space has been

²Throughout this paper, I use typewriter font family exclusively for the name of packages available in R. For instance, **BMA** refers to the name of a package and BMA refers to the Bayesian model averaging technique.

set the posterior distribution of the parameters of interest from a linear model (say), β , given the data, Δ , is

$$Pr(\beta|\Delta) = \sum_{j=1}^K Pr(\beta|M_j, \Delta) Pr(M_j|\Delta), \quad (2.1)$$

which is simply the weighted average of the posterior distributions under each model. Using Bayes Theorem, one can rewrite the posterior probability of model M_k as follows

$$Pr(M_k|\Delta) = Pr(\Delta|M_k) \frac{Pr(M_k)}{Pr(\Delta)} = Pr(\Delta|M_k) \frac{Pr(M_k)}{\sum_{j=1}^K Pr(\Delta|M_j) Pr(M_j)} \quad (2.2)$$

$$Pr(\Delta|M_k) = \int Pr(\Delta|\beta^{(k)}, M_k) Pr(\beta^{(k)}|M_k) d\beta^{(k)} \quad (2.3)$$

where $\beta^{(k)}$ is the vector of parameters from model M_k , $Pr(\beta^{(k)}|M_k)$ is a prior probability distribution assigned to the parameters of model M_k , and $Pr(M_k)$ is the prior probability that M_k is the true model. As mentioned above, in order to make BMA applicable, I need to choose prior distributions for the parameters of the model and then, based on the data, calculate posterior probabilities attached to different models and finally, find the parameter distributions by averaging the posterior distributions.

One should notice that, as stated in Brock et al. (2003), the set of prior probabilities must be informative with respect to the likelihood, meaning priors should be relatively high where the likelihoods are large otherwise the choice of priors will have a substantial effect on the posteriors. Several studies have attempted to discern the effect that the choice of prior probabilities has on posterior probabilities, which are commonly used to gauge the ‘robustness’ of a variable. Ley & Steel (2009a) provide a complete review on the effect of prior assumptions on the outcome of BMA analyses. Additionally, Eicher et al. (2011) address the issue of which default prior to use in applied settings when information on the prior is lacking or small relative to the information in the data. They use cross-validated predictive performance and find support for the unit information prior (UIP) for the integrated likelihood coupled with a uniform prior over the model space. The key implication of their work is that the UIP is the safest and most robust choice to use as a benchmark. However, it is not clear that a specific prior will always emerge as the best prior to use across all econometric settings.

2.3.2 Weighted Average Least Squares

Magnus et al. (2010) recently introduced the WALS estimator based on the Equivalence Theorem of Mean Squared Error (MSE), discussed in Magnus & Durbin (1999) and Danilov & Magnus (2004). The WALS estimator, which is a Bayesian combination of Frequentist estimators, takes the required prior probabilities from the Laplace distribution and generates bounded risk, unlike standard BMA estimators (which typically deploy normal priors). To

implement the WALS estimator, assume a linear regression model

$$y = X\beta + \varepsilon = X_1\beta_1 + X_2\beta_2 + \varepsilon, \quad \varepsilon \sim i.i.d N(0, \sigma^2) \quad (2.4)$$

where y is an $n \times 1$ vector of the dependent variable, (β_1, β_2) is the vector of model parameters, X_1 ($n \times k_1$), $X_1 \neq \emptyset$, contains variables that will always be included in the model (such as an intercept), X_2 ($n \times k_2$) includes auxiliary variables which may be included in the model and $X = (X_1 : X_2)$, the matrix of regressors, has full column-rank.³

Initially I orthogonalize the columns of X_2 such that $Q'X_2'H_1X_2Q = \Lambda$, where $H_1 = I_n - X_1(X_1'X_1)^{-1}X_1'$, Λ is a diagonal matrix, and Q is an orthogonal $k_2 \times k_2$ matrix. Consequently,

$$X_2\beta_2 = X_2Q\Lambda^{-1/2}\Lambda^{1/2}Q'\beta_2 = \tilde{X}_2\beta_2^* \quad (2.5)$$

This suggests that I can replace $X_2\beta_2$ in equation (2.4) with $\tilde{X}_2\beta_2^*$, estimate (β_1, β_2^*) rather than (β_1, β_2) , and then recover β_2 from $\beta_2 = Q\Lambda^{-1/2}\beta_2^*$. Given that any combination of the columns of X_2 can be used to evaluate (2.4), there are a total of $K = 2^{k_2}$ models to be evaluated. For model j , $k_{2j} \leq k_2$ columns of X_2 are selected. The estimators of β_1 and β_2 under model j , after some simple algebra, can be written as

$$\hat{\beta}_1^{(j)} = (X_1'X_1)^{-1}X_1'y - (X_1'X_1)^{-1}X_1'X_2W_jX_2'H_1y, \quad \hat{\beta}_2^{(j)} = W_jX_2'H_1y$$

where W_j is a diagonal $k_2 \times k_2$ matrix with k_{2j} ones and $k_2 - k_{2j}$ zeros on the diagonal, such that the m th diagonal element is zero if the m th element of $\beta_2^{(j)}$ is zero. The WALS estimator of β_1 is simply a weighted average of the $\hat{\beta}_1^{(j)}$ s

$$\hat{\beta}_{1,WALS} = \sum_{j=1}^{2^{k_2}} \lambda_j \hat{\beta}_1^{(j)} = (X_1'X_1)^{-1}X_1'y - (X_1'X_1)^{-1}X_1'X_2W X_2'H_1y$$

where $W = \sum_j \lambda_j W_j$. The Equivalence Theorem states that

$$E(\hat{\beta}_{1,WALS}) = \beta_1 - (X_1'X_1)^{-1}X_1'X_2E(W\hat{\beta}_2 - \beta_2), \quad \text{where} \quad \hat{\beta}_2 = X_2'H_1y$$

This means that $\hat{\beta}_{1,WALS}$ is a “good” estimator of β_1 if and only if $W\hat{\beta}_2$ is a “good” estimator of β_2 . $W = \text{diag}(\omega_1, \omega_2, \omega_3, \dots, \omega_{k_2})$ is a diagonal matrix and the components of $\hat{\beta}_2$ are independent because of the model space orthogonalization. Therefore, if I choose ω_j to be a function of $\hat{\beta}_{2j}$ (a scalar) only, then the components of $W\hat{\beta}_2$ will be independent. Namely, my 2^{k_2} -dimensional problem reduces to k_2 one-dimensional problems. The importance of this result is that the WALS estimator, unlike the BMA estimator, can fully enumerate the model space even for large k_2 since it is *linear* in the number of covariates under inspection. A standard BMA implementation is exponential in the number of covariates and even for a moderate number of covariates MCMC methods are used to sample from the model space since full enumeration is computationally expensive (if not impossible).

³The error terms are assumed to be distributed identically and independently while a more general assumption on the error term for the WALS estimator is discussed in Magnus, Wan & Zhang (2011).

Using the fact that $\hat{\beta}_{2j}|\beta_{2j} \sim N(\beta_{2j}, \sigma^2)$, the Laplace estimator of β_{2j} is defined as

$$\tilde{\beta}_{2j} = E(\beta_{2j}|\hat{\beta}_{2j}) = \frac{1 + \delta(\hat{\beta}_{2j})}{2}(\hat{\beta}_{2j} - c) + \frac{1 - \delta(\hat{\beta}_{2j})}{2}(\hat{\beta}_{2j} + c)$$

where

$$\delta(\hat{\beta}_{2j}) = \frac{e^{-c\hat{\beta}_{2j}}\Phi(\hat{\beta}_{2j} - c) - e^{c\hat{\beta}_{2j}}\Phi(-\hat{\beta}_{2j} - c)}{e^{-c\hat{\beta}_{2j}}\Phi(\hat{\beta}_{2j} - c) + e^{c\hat{\beta}_{2j}}\Phi(-\hat{\beta}_{2j} - c)}$$

$\Phi(\cdot)$ represents the CDF of the standard normal distribution, and c is a positive hyperparameter that usually is set to be $c = \log(2)$ for the Laplace distribution. Define $\tilde{\beta}_{2,Laplace} = (\tilde{\beta}_{21}, \tilde{\beta}_{22}, \tilde{\beta}_{23}, \dots, \tilde{\beta}_{2k_2})$. Then the WALS estimators for β_1 and β_2 are

$$\hat{\beta}_{2,WALS} = \sigma Q \Lambda^{-1/2} \tilde{\beta}_{2,Laplace} \quad (2.6)$$

$$\hat{\beta}_{1,WALS} = (X_1' X_1)^{-1} X_1' (y - X_2 \hat{\beta}_{2,WALS}) \quad (2.7)$$

See Magnus et al. (2010, Section 3.5) for further details on implementation of the WALS estimator.⁴

2.3.3 Frequentist Model Averaging

An alternative method to deal with model uncertainty is FMA. Unlike model selection approaches where the researcher has to select one model, according to a criterion, from the set of acceptable models $M = \{M_1, M_2, \dots, M_K\}$, and unlike BMA, which eventually computes the probability of being in the final model for all available variables, an FMA estimator assigns weights, based upon a specific criterion, to each model, M_j , $j = 1, 2, \dots, K$, and then takes the weighted average of all the parameter vectors to obtain the model average estimator. Mathematically, the estimator of the parameters of interest, β , is

$$\hat{\beta} = \sum_{j=1}^K \omega_j \hat{\beta}^{(j)} \quad (2.8)$$

where ω_j , $j = 1, 2, \dots, K$ are weights normalized such that $\sum_j \omega_j = 1$, and assigned to the corresponding models, and $\hat{\beta}^{(j)}$ is the estimate of β under model j . Buckland et al. (1997) suggest calculating the weight vector based on an exponential AIC criterion while Hansen (2007) suggests using a Mallows's criterion for the selection of the weights. The use of the Mallows's criterion is desirable since both Hansen (2007) and Wang et al. (2009) show that asymptotically the model average estimator in (2.8) achieves the lowest possible squared error in the class of model average estimators.⁵

To implement the MMA approach laid out in Hansen (2007) consider a standard linear

⁴See also Magnus et al. (2011) for a formal application.

⁵See Li (1987) for a similar proof in the model selection context.

regression model

$$y = X\beta + \varepsilon = \sum_{i=1}^{\infty} \beta_i x_i + \varepsilon, \quad \varepsilon \sim i.i.d \ D(0, \sigma^2) \quad (2.9)$$

where y is the $n \times 1$ vector of dependent observations and β is the vector of model parameters. Let X_{k_j} be the corresponding submatrix of X containing the first k_j columns of X .⁶ Index $j = 1, 2, 3, \dots, K$ indicates the sequence of models. I estimate the ℓ th model, $y = X_{k_\ell} \beta^{(\ell)} + \varepsilon_\ell$, by least squares and denote the vector of estimated coefficients under model ℓ as $\hat{\beta}^{(\ell)}$ and the vector of estimated residuals as $\hat{\varepsilon}_\ell$.

This procedure is repeated for all K models and the final estimator of β is presented in equation (2.8), where ω_j , $j = 1, 2, \dots, K$ are normalized weights calculated numerically by minimizing the Mallows criterion,

$$C_p(\Omega) = \Omega' \hat{\varepsilon}' \hat{\varepsilon} \Omega + 2\sigma^2 L' \Omega \quad (2.10)$$

with $\Omega_{(K \times 1)} = (\omega_1, \omega_2, \dots, \omega_K)'$ representing the vector of weights, $\hat{\varepsilon}_{(n \times K)} = (\hat{\varepsilon}_1, \hat{\varepsilon}_2, \dots, \hat{\varepsilon}_K)$ is the matrix of all residual vectors, and $L_{(K \times 1)} = (k_1, k_2, \dots, k_K)'$ is the vector of the number of regressors used for each of the K models.⁷ This is, in fact, a classic linear-quadratic programming problem. Mathematically, the optimization problem is

$$\begin{aligned} \min_{\omega_j} C_p(\Omega) &= \Omega' \hat{\varepsilon}' \hat{\varepsilon} \Omega + 2\sigma^2 L' \Omega \\ \text{s.t.} \quad \sum_{j=1}^K \omega_j &= 1, \quad 0 \leq \omega_j \leq 1, \quad \forall j = 1, 2, \dots, K. \end{aligned} \quad (2.11)$$

A solution for this minimization problem can be found in any number of programming languages.⁸ One can replace the unknown σ^2 with the estimator $\hat{\sigma}_K^2 = (n - K)^{-1} \hat{\varepsilon}_K' \hat{\varepsilon}_K$. Hansen (2007, Theorem 2) shows that $\hat{\sigma}_K^2$ is a consistent estimator for σ^2 provided $K/n \rightarrow 0$ and $K \rightarrow \infty$ as the sample size increases. Moreover, replacing σ^2 with an estimator does not affect the asymptotic optimality properties of the weight vector selected via the Mallows criterion.

Unfortunately, no formal asymptotic theory exists for the MMA estimator to obtain the variance-covariance matrix needed to construct standard errors. I therefore employ Buckland et al.'s (1997) suggestion to calculate the variance of my MMA estimator based on the bootstrap. Their conservative approach finds the maximum possible value for variances that incorporate a component representing misspecification bias. Buckland et al. (1997) and Schomaker, Wan & Heumann (2010) propose estimating the variance of an FMA estimator as

$$\text{var}(\hat{\beta}) = \sum_{j=1}^K \omega_j \sqrt{\text{var}(\hat{\beta}^{(j)} | \text{bias}_j) + \text{bias}_j^2} \quad (2.12)$$

⁶Since I assume $E(\varepsilon|X) = 0$ the first column of X can be thought of as a vector of ones representing the intercept.

⁷A thorough review of MMA is provided in Wang et al. (2009).

⁸I use the **quadprog** package (Turlach & Weingessel 2007) available in R.

that accounts for the additional uncertainty due to model selection. In this formula $bias_j = \hat{\beta}^{(j)} - \hat{\beta}$ is the misspecification bias that arises in estimating β under model j . I use the full model, including all covariates, as the benchmark to calculate the misspecification bias. That is, I use OLS to estimate the vector of parameters, $\beta^{(j)}$, under model j , and also the variances of the estimated parameters. Then I subtract the vector of estimated parameters under model j from the full model to construct the bias for model j . As Hjort & Claeskens (2003) mentioned, this formula gives reasonably accurate standard errors that include additional uncertainty due to model selection.

2.3.3.1 Enumerating the Model Space

To implement the MMA estimator two key issues remain. First, Hansen (2007, Theorem 1) requires an explicit ordering of the regressors. Fortunately, Wan et al. (2010) established the optimality of MMA without requiring the explicit ordering needed in Hansen's (2007) proofs.⁹ Second, ignoring the ordering, the question of how best to select variables and combine them for all the models to be weighted within the Mallows's criterion is paramount. Here, I use the pioneering insight of Magnus et al. (2010) to reduce the total size of the model space from $K = 2^k - 1$ to $K = k$ where k is the largest number of covariates that one has. This result is obtained by showing that the key results of Hansen do not change when I have an orthogonalization of the entire set of covariates.

For the moment I will assume that σ^2 is known. Use the notation X_m to denote the matrix of covariates used to construct the OLS estimates for the m th model. The OLS estimator for the m th model is $\hat{\beta}^{(m)} = (X_m' X_m)^{-1} X_m' y$. The matrix $P_m = X_m (X_m' X_m)^{-1} X_m'$ will figure prominently in the following discussion. What I want is to have X_m such that

$$Q' X_m' X_m Q = \Lambda,$$

where Q is an orthogonal matrix and Λ is a diagonal matrix. I then have

$$\Lambda^{-1/2} Q' X_m' X_m Q \Lambda^{-1/2} = I$$

Letting $\tilde{X}_m = X_m Q \Lambda^{-1/2}$ I have $\tilde{X}_m' \tilde{X}_m = I$. I note that Lemma 1 in Hansen (2007) still holds because

$$\tilde{P}_m = \tilde{X}_m (\tilde{X}_m' \tilde{X}_m)^{-1} \tilde{X}_m' = \tilde{X}_m \tilde{X}_m' = X_m Q \Lambda^{-1} Q' X_m' = X_m (X_m' X_m)^{-1} X_m' = P_m.$$

Also, it is trivial to recover $\beta^{(m)}$ from $\beta^{*(m)}$ (the coefficient vector from the orthogonalization) since

$$\tilde{X}_m \beta^{*(m)} = \tilde{P}_m y = P_m y = X_m \beta^{(m)}.$$

This yields $\beta^{*(m)} = \Lambda^{1/2} Q' \beta^{(m)}$ so once I have estimates of $\beta^{*(m)}$ I can recover estimates of

⁹In results not reported here for my replication studies, I do indeed obtain approximately identical coefficient estimates when I consider different orderings of the regressors within my MMA framework. Furthermore, see Wan et al. (2010, Example 2) which shows how considering a set of nested models satisfies the requisite conditions for the non-ordering to hold, provided the model space grows slow enough.

the parameters of interest as $\beta^{(m)} = Q\Lambda^{-1/2}\beta^{*(m)}$.

The importance of the orthogonalization is that the residuals, $\hat{\varepsilon}_m = H_m y = (I - P_m)y$ are identical across the K models that I estimate for the model averaging. The independence amongst the covariates means that when a new variable is added to the model, the information it contains is not corrupted by covariances with the other variables, implying that no additional combinations of the variables are needed to exhaust the model space. In effect, what the orthogonalization gives me is the ability to select a single variable and continue to add to the model space *linearly*. Also, given that the variables are orthogonal the order in which this is done is immaterial. The simplest way to orthogonalize the vector space is to use the singular value decomposition or Gram-Schmidt orthogonalization.

Once the orthogonalization has been undertaken the MMA weights can be found by simply estimating a set of models which increase by a single variable each time.¹⁰ For example, suppose that there were 5 regressors which one wanted to average over. The total number of models to average over would be $2^5 - 1 = 31$. However, if I orthogonalize the regressors then I have only 5 models to average over $\{1, (1, 2), (1, 2, 3), (1, 2, 3, 4), (1, 2, 3, 4, 5)\}$. This may seem trivial for a small number of regressors, but it quickly becomes computationally relevant. With just 20 regressors, not uncommon in applied work, I already have more than one million elements of the model space. It is clear that both MMA and WALs can be implemented across a large number of covariates without concern that enough models are visited, as in BMA investigations.

2.4 Data

My basis for comparison of BMA, WALs and MMA techniques are three recent studies of growth theory uncertainty. I first replicate the results of Fernandez et al. (2001*b*), Masanjala & Papageorgiou (2008) and Doppelhofer & Weeks (2009) using freely available software that conducts BMA. Second, I compare the BMA posterior means and standard errors to those obtained using both the MMA and WALs estimators described above for the set of variables deemed important in each paper. All three datasets are publicly available on the Journal of Applied Econometrics online data archive. Due to the fact that these datasets are well documented I briefly discuss them here and direct the reader to more detailed treatments.

My first study, Fernandez et al. (2001*b*) (FLS hereafter) uses a cross section of 72 countries along with 41 potential growth determinants for the period 1960 to 1992. This dataset is taken from the larger dataset used by Sala-i-Martin (1997) for his study on robust determinants of growth. My second dataset is a subset of the FLS data used by Masanjala & Papageorgiou (2008) (MP hereafter). MP's focus was on growth theory uncertainty and if the determinants of growth differed for African countries relative to the rest of the world. They used the 37 sub-Saharan African countries within the FLS dataset along with 24 of the 41 variables. See the appendix in MP for the countries and variables used in this study. My last dataset comes from Doppelhofer & Weeks (2009) (DW hereafter) who used the data in Sala-i-Martin, Doppelhofer & Miller (2004) to study jointness of growth determinants.

¹⁰Given the independence of the regressors the order in which this is done is irrelevant.

This data set comprises 88 countries and 67 candidate variables in a cross section for the period 1960 to 1996. All 67 of the variables used can be found in data appendix B in DW.

2.5 Results

To implement the BMA methods in the three papers of interest I first consider the similarity of the results from the 3 available packages currently available in *R* (BMA, BAS, and BMS). Amini & Parmeter (2011) detail the features of each of these three packages and advocate on behalf of the BMS package. Table 2.1 presents the posterior inclusion probabilities (PIP), posterior means and standard deviations from each of the packages using the Fernandez et al. (2001b) study as a benchmark. For succinctness, in the remainder of the paper I only present results for variables with PIPs greater than 0.5 from the published studies. As is apparent, only the BMS package is successful at matching the reported PIPs and posterior mean/standard deviation while the BAS package comes close to reproducing the posterior means/standard deviations but the PIPs display significant differences (compare the results for Rule of Law and Years Open Economy). The BMA package also produces similar posterior means but cannot achieve the same accuracy of the PIPs. Given these results, coupled with the recommendation of Amini & Parmeter (2011) I present all remaining BMA results using the BMS package.¹¹

In passing I note that the results in Table 2.1 are not due to the specific study under comparison. For all three of the studies considered here the BAS package cannot replicate the full set of estimates (PIPs, posterior mean/standard deviations) while the BMA package can only reproduce the results of the Masanjala & Papageorgiou (2008) study, primarily because this study used the BMA package for estimation.¹² My intuition for this breakdown between the BMA, BMS and BAS packages is that the BMA package does not use a Monte Carlo Markov Chain (MCMC) algorithm to search over the model space¹³ and the available options within the packages does not allow the priors to be identical across the three packages; see Amini & Parmeter (2011) for more details on the options available within each of these packages.

For the following analyses I construct BMA, WALs, and MMA estimators using the three datasets, FLS, MP, and DW, and compare the findings. Implementing the BMA estimator requires specifying the set of model priors and a search algorithm. I use the uniform distribution for model priors¹⁴ and a MCMC algorithm to search over the model space to construct my BMA estimator. I also use the *g*-prior framework, as introduced in Fernandez

¹¹PIPs and posterior mean and standard deviations stemming from the additional packages are available upon request.

¹²A working paper version of Masanjala & Papageorgiou (2008) possesses different results than the published version which, coincidentally, the BMA package cannot fully replicate while the BMS package can.

¹³The BMA package evaluates the most visited models based upon a Bayesian Information Criterion difference and a hierarchical OLS *t*-test that is different from the methods used in both the BMS and BAS packages.

¹⁴Note that the imposition of a uniform prior across the model space is not innocuous, and can materially impact inference. A particular problem in adopting such an uninformative uniform prior is that the expected prior model size can be very large in cases where K , the number of covariates, is large. In other words the great majority of prior probability is allocated to models with a large number of variables.

Table 2.1: Performance of the BMS, BAS, and BMA Packages for the Fernandez et al. (2001*b*) Study.

	Published Results	BMS Package		BAS Package		BMA Package	
	PIP (i)	PIP (ii)	Post Mean & SD (iii)	PIP (iv)	Post Mean & SD (v)	PIP (vi)	Post Mean & SD (vii)
Variable:							
GDP 1960	1.00	1.00	-0.0160 (0.0031)	1.00	-0.0168 (0.0026)	1.00	-0.0173 (0.0026)
Fraction Confucian	0.995	0.990	0.0563 (0.0147)	1.00	0.0667 (0.0115)	1.00	0.0590 (0.0109)
Life Expectancy	0.946	0.941	0.0008 (0.0003)	0.999	0.0008 (0.0002)	1.00	0.0008 (0.0002)
Equipment Investment	0.942	0.939	0.1594 (0.0685)	0.985	0.1373 (0.0446)	0.972	0.0993 (0.0440)
Sub-Saharan Dummy	0.757	0.751	-0.0118 (0.0084)	0.996	-0.0211 (0.0051)	1.00	-0.0184 (0.0053)
Fraction Muslim	0.656	0.634	0.0084 (0.0077)	0.430	0.0049 (0.0068)	0.519	0.0055 (0.0061)
Rule of Law	0.516	0.510	0.0074 (0.0082)	0.942	0.0110 (0.0048)	1.00	0.0126 (0.0043)
Years Open Economy	0.502	0.500	0.0069 (0.0079)	0.049	0.0001 (0.0014)	0.055	0.0003 (0.0018)

Notes: This table presents the results of applying Bayesian model averaging to the data used in Fernandez et al. (2001*b*) using the BMS, BAS, and BMA packages. The dependent variable is the growth rate from 1960-1992 across 72 countries. Standard errors are in parenthesis. All models estimated contain an intercept. Column (i) shows the Bayesian posterior inclusion probabilities that Fernandez et al. (2001*b*) have reported. Columns (ii), (iv), and (vi) report the PIPs using BMS, BAS, and BMA packages, respectively. Columns (iii), (v), and (vii) indicate BMA posterior estimates of the model coefficients and standard errors using the three packages.

et al. (2001*a*), as the benchmark for prior assumptions of the model parameters. My implementation of the WALS estimator begins by orthogonalizing the entire dataset and then using the moments of the Laplace distribution explained in section 2.3.2 to compute the appropriate weights using $c = \log(2)$ as the hyperparameter. I use a similar methodology to implement the MMA estimator. I add one variable linearly at a time to the set of covariates, orthogonalize the vector space constructed by those variables, and then compute the weights by minimizing the Mallows's criterion provided in (2.11). The total number of models visited by both WALS and MMA is k , the number of covariates in each dataset, given my discussions in sections 2.3.2 and 2.3.3.1. Regardless of model averaging estimator (BMA, WALS, MMA) all models estimated contain an intercept.

2.5.1 A comparison using Fernandez, Ley & Steel (2001*b*)

Table 2.2 presents the BMA, WALS and MMA estimates for the Fernandez et al. (2001*b*) study for the variables with PIPs above 50%. Column (i) presents the PIPs, calculated via the BMS package. The explanatory variables are sorted from the highest PIP to the lowest. Column (ii) reports the estimated posterior means and their standard errors. Column (iii) reports the model average estimates and standard deviations using the WALS estimator of Magnus et al. (2010) while column (iv) reports the corresponding model average estimates using the MMA estimator of Hansen (2007) using the orthogonalization described earlier.

Comparing across the three set of estimates I highlight that both the sign and magnitude of all three sets of estimates are identical. However, the three estimators deliver different point estimates and standard deviations for each of the variables. While not a strict relationship, the MMA point estimates appear to be marginally larger in absolute value for all seven variables relative to the WALS estimates (excluding Fraction Muslim). A strict relationship between the three sets of estimated standard errors does not appear to exist.

One important highlight from the comparison of the three estimation approaches here is that the results are relatively ‘close’ and part of this in term stems from the fact that all three frameworks use a uniform prior on the model space. The ability to place differing probability mass on different models throughout the model space is a strength of the Bayesian approach. While the researcher can estimate, or feed in specific models to the WALS and MMA setups, the full generality of the model prior is an advantage of using BMA. This is an important issue that, while beyond the scope of this paper, certainly requires further study within the FMA estimation paradigm.¹⁵

2.5.2 A comparison using Masanjala & Papageorgiou (2008)

Masanjala & Papageorgiou (2008) study the sources of growth in Africa. This was an important study because not only did it address model uncertainty, but it also focused on parameter heterogeneity. Their results find key determinants of growth in Africa that are substantially different from those in the rest of the world. They conclude that mining, primary exports, and primary school enrollment are the main drivers behind sub-Saharan African growth. The published version of the paper produces results using the BMA package and as such replication of this paper is trivial.¹⁶

Table 2.3 reports the same set of comparison results as Table 2.2 for Masanjala & Papageorgiou’s (2008) study, using those variables with estimated PIPs higher than 50%. Column (i) presents the PIPs. The explanatory variables are sorted from the highest PIP to the lowest. Column (ii) reports the estimated posterior means and their standard errors. Column (iii) reports the model average estimates and standard deviations using WALS while column (iv) reports the corresponding model average estimates using the MMA estimator with the variable orthogonalization described earlier. I mention again that replication of Masanjala

¹⁵I reexamined the PIPs and posterior means from my BMA results using two alternative sets of priors. First, I used “fixed” priors corresponding to fixed common prior inclusion probabilities for each regressor, as described in Sala-i-Martin et al. (2004). In this case the PIPs and posterior means are almost identical to the current case employing uniform priors. Second, I used “random” priors that trigger the ‘random theta’ prior by Ley & Steel (2009a), who suggest a binomial-beta hyperprior for the prior inclusion probability. In this case the PIPs and posterior means are significantly different from the other two settings. Overall, it seems that, as stated in Brock et al. (2003), results stemming from a BMA analysis are sensitive to the choice of model priors. Additionally, the fact that the posterior means differ substantially also suggests that examining MMA against a nonuniform prior BMA approach may not necessarily represent a valid comparison.

¹⁶However, the BMA package was unable to replicate the results of the working paper version of MP whereas the BMS package was able to obtain identical estimates. Moreover, the variables that are relevant are different between these two versions of the paper suggesting that differences in model priors and search algorithms necessary for estimation can produce incongruous insights.

Table 2.2: BMA, WALs, and MMA Estimation for Fernandez et al. (2001b) Study.

Method:	PIP-BMA	BMA	WALS	MMA
	(i)	(ii)	(iii)	(iv)
Variable:				
GDP 1960	1.00	-0.0160 (0.0031)	-0.0139 (0.0029)	-0.0169 (0.0032)
Fraction Confucian	0.990	0.0563 (0.0147)	0.0663 (0.0141)	0.0747 (0.0150)
Life Expectancy	0.941	0.0008 (0.0003)	0.0007 (0.0002)	0.0008 (0.0002)
Equipment Investment	0.939	0.1594 (0.0685)	0.1041 (0.0378)	0.1272 (0.0445)
Sub-Saharan Dummy	0.751	-0.0118 (0.0084)	-0.0157 (0.0061)	-0.0201 (0.0067)
Fraction Muslim	0.634	0.0084 (0.0077)	0.0108 (0.0079)	0.0106 (0.0086)
Rule of Law	0.510	0.0074 (0.0082)	0.0089 (0.0058)	0.0116 (0.0061)

Notes: This table presents the Bayesian model averaging (BMA), Weighted average least squares (WALS), and Mallows model averaging (MMA) techniques applied to the data used in Fernandez et al. (2001b). The dependent variable is the growth rate from 1960-1992 across 72 countries. Standard errors are in parenthesis. All models estimated contain an intercept. Column (i) reports the Bayesian posterior inclusion probabilities (PIPs), calculated via the BMS package. Column (ii) reports the Bayesian posterior means and their standard errors. Column (iii) reports the estimation of model coefficients and standard deviations using the WALS approach discussed in Magnus et al. (2010). Column (iv) reports the corresponding model averaging estimates from the MMA estimator of Hansen (2007) using the orthogonalization of model space. The BMA analysis was set up with a uniform prior on the model space.

& Papageorgiou (2008) is trivial given that they used the BMA package.

A key difference that arises between the BMA point estimates and those of WALS and MMA is the variables Years Open Economy and Mining. First, Years Open Economy has a positive effect using BMA while it has a negative effect using either WALS or MMA. Second, Mining has a positive effect for all three model averaging methods, yet, the magnitude of the WALS point estimate is almost three times larger than that of BMA while the MMA results are more than three times larger than the posterior means from BMA. Even more important, both WALS and MMA suggest that Mining is statistically relevant, as judged by the t -ratio while the BMA results suggest statistical insignificance. Contrary to what I witnessed in Table 2.2, here the three competing methods yield different insights.

2.5.3 A comparison using Doppelhofer & Weeks (2009)

Doppelhofer & Weeks (2009) recently focused on model averaging exercises where jointness was considered. Jointness of explanatory variables is based on the joint posterior distribution of variables over the model space as opposed to marginal posterior distributions, which

Table 2.3: BMA, WALS and MMA Estimation for Masanjala & Papageorgiou (2008) Study.

Method:	PIP-BMA	BMA	WALS	MMA
	(i)	(ii)	(iii)	(iv)
Variable:				
GDP 1960	0.824	-0.0106 (0.0060)	-0.0162 (0.0041)	-0.0195 (0.0041)
Years Open Economy	0.732	0.0309 (0.0215)	-0.0264 (0.0275)	-0.0367 (0.0317)
Mining	0.670	0.0398 (0.0328)	0.0928 (0.0305)	0.1242 (0.0348)
Primary Exports	0.538	-0.0208 (0.0221)	-0.0160 (0.0165)	-0.0258 (0.0187)
Primary School Enrollment	0.514	0.0005 (0.0006)	-0.0001 (0.0005)	-0.0001 (0.0005)

Notes: This table presents the Bayesian model averaging (BMA), Weighted average least squares (WALS), and Mallows model averaging (MMA) techniques applied to the data used in Masanjala & Papageorgiou (2008). The dependent variable is the growth rate from 1960-1992 across 37 sub-Saharan African countries. Standard errors are in parenthesis. All models estimated contain an intercept. Column (i) reports the Bayesian posterior inclusion probabilities (PIPs), calculated via the **BMS** package. Column (ii) reports the Bayesian posterior means and their standard errors. Column (iii) reports the estimation of model coefficients and standard deviations using the WALS approach discussed in Magnus et al. (2010). Column (iv) reports the corresponding model averaging estimates from the MMA estimator of Hansen (2007) using the orthogonalization of model space. The BMA analysis was set up with a uniform prior on the model space.

were the focus of the previous two studies. Moreover, jointness is an important concept to capture because it impacts statistical inference and can be informative for economic policy. Doppelhofer & Weeks (2009, Appendix B) provides the BMA PIPs, posterior means and standard errors. Table 2.4 presents the estimates for the three different model averaging methods. The BMA results are exactly replicated from the Doppelhofer & Weeks (2009) study. Similar to what I witnessed from Table 2.2, the results from the BMA, WALS and the MMA methods all produce coefficient estimates which are similar in sign and magnitude.

A consistent theme arising from all three studies is that the standard errors from MMA are slightly larger than those from WALS and both methods have standard errors roughly double than those found using BMA. One possible explanation for this is that the construction of the variances of the coefficient estimates for the MMA estimator are conservative given that it is not based on a formal asymptotic representation. An alternative explanation, at least for the DW study is that for the implementation of the MCMC model search, the **BMS** package only needs to visit 29,000 models out of 2^{68} models to achieve convergence. Compare this with the FLS dataset, where the **BMS** package needs to visit 52,000 models out of 2^{42} to achieve convergence, which in percentage terms is more than 100 million times larger than the percentage of potential models visited in the DW dataset using the same MCMC model search algorithm. In this sense the BMA results for the DW study are highly accurate, using models with very high likelihood. Moreover, this fact coupled with the full enumeration of the model space by WALS and MMA means that many poor models could

have been used which adds to the uncertainty of the average coefficient estimates, resulting in larger standard errors.

Ignoring the root cause of the discrepancies in the standard errors across the three methods, even with the conservative nature of the MMA standard errors, the significance of the weighted estimates for MMA across the three examples is roughly identical to those stemming from WALS and are broadly consistent with the BMA results. A formal derivation of the limiting distribution of the MMA estimator marks an interesting avenue for future research.

Table 2.4: BMA, WALS, MMA Estimation for Doppelhofer & Weeks (2009) Study.

Method:	PIP-BMA	BMA	WALS	MMA
	(i)	(ii)	(iii)	(iv)
Variable:				
East Asian Dummy	0.848	0.0182 (0.0094)	0.0116 (0.0155)	0.0132 (0.0193)
Primary Schooling Enrollment	0.798	0.0215 (0.0128)	0.0204 (0.0163)	0.0304 (0.0205)
Investment Price	0.790	-0.00006 (0.00003)	-0.00008 (0.00004)	-0.0001 (0.00005)
Initial GDP 1960	0.742	-0.0062 (0.0045)	-0.0053 (0.0072)	-0.0081 (0.0091)
Fraction of Tropical Area	0.591	-0.0085 (0.0078)	-0.0090 (0.0154)	-0.0144 (0.0193)

Notes: This table presents the Bayesian model averaging (BMA), Weighted average least squares (WALS), and Mallows model averaging (MMA) techniques applied to the data used in Doppelhofer & Weeks (2009). The dependent variable is the growth rate from 1960-1996 across 88 countries. Standard errors are in parenthesis. All models estimated contain an intercept. Column (i) reports the Bayesian posterior inclusion probabilities (PIPs), calculated via the BMS package. Column (ii) reports the Bayesian posterior means and their standard errors. Column (iii) reports the estimation of model coefficients and standard deviations using the WALS approach discussed in Magnus et al. (2010). Column (iv) reports the corresponding model averaging estimates from the MMA estimator of Hansen (2007) using the orthogonalization of model space. The BMA analysis was set up with a uniform prior on the model space.

2.6 Jackknife Model Averaging

The optimality of MMA fails under heteroscedastic error structure. A recently proposed estimator, Jackknife model averaging (JMA), that deals with parameter estimation under model uncertainty in *heteroscedastic* error settings, has major advantages over other existing model selection or Frequentist model averaging techniques. This estimator, which minimizes a cross-validation (leave-one-out) criterion to obtain the appropriate weights for averaging all the candidate models, asymptotically achieves the lowest possible expected squared errors. In practice, however, considering all the combinations of covariates is computationally expensive or infeasible in the presence of large number of potential determinants. In this section, I show that model space orthogonalization drastically reduces the size of space over

which I perform model selection and enables me to select a single variable and continue to add to the model space linearly rather than exponentially. As an illustration, the modified JMA technique is applied to a housing dataset, and the results are compared to an alternative model averaging method.

In addition to MMA and FIC estimators, Hansen & Racine (2009) introduced another estimator within the Frequentist model averaging framework called Jackknife model averaging (JMA) that selects appropriate weights for averaging models by minimizing a cross-validation criterion. All these methods encapsulate model averaging within the Frequentist econometric arena and are directly comparable to BMA tools. In practice, the total number of possible model candidates is 2^k where k indicates the number of covariates. Consider a researcher that has collected data on 30 potential variables that she thinks they belong to the data generating process with a certain probability. In the absence of a comprehensive theoretical framework, she has to deal with more than a billion different models (precisely 2^{30}) to assess the true size of each variable's influence on the dependent variable. Using any of the proposed criteria to calculate the weights for this size of model space is computationally near to infeasible.

In this section, I seek to address this issue using the insights of Magnus et al.'s (2010) work. I show that if I orthogonalize the model space prior to the process of computing the weights, the size of space over which I perform model selection will drop from 2^k to only k models. In other words, this orthogonalization enables me to select a single variable and continue to add to the model space linearly rather than exponentially. My focus is on applying model space orthogonalization to JMA estimator that has desirable optimality properties. It allows for heteroscedastic error terms and reaches the lowest possible expected squared error over the class of linear estimators. I demonstrate JMA performance, in an orthogonalized model space, through an empirical example where I use a hedonic price dataset and directly compare the JMA results with similar analysis using BMA. I find that my final weighted estimated parameter vectors are nearly identical to the posterior means obtained using BMA.

2.7 Jackknife Model Averaging Estimator

Hansen & Racine (2009) proposed an estimator within the Frequentist model averaging framework named Jackknife Model Averaging (JMA) that, unlike the MMA, is able to estimate a conditional mean optimally in the face of model uncertainty in heteroscedastic error setting. This estimator selects the weights of the different models by minimizing a cross-validation (leave-one-out) criterion and is asymptotically optimal in the sense of reaching the lowest possible squared errors over the class of linear estimators. Implementation of JMA estimator is as follows.

Suppose that a sample of n observations $(y_{i(1 \times 1)}, x_{i(1 \times k)}), i = 1, 2, \dots, n$, is given and one believes that data generating process is $y_i = \mu_i + e_i$, where μ_i is the conditional mean of y_i given x_i , and e_i is a heteroscedastic error term such that $E(e_i|x_i) = 0$ and $E(e_i^2|x_i) = \sigma_i^2$. Let define the vector of observations on independent variable as $y_{n \times 1} = (y_1, y_2, \dots, y_n)'$,

the vector of fitted values as $\mu_{n \times 1} = (\mu_1, \mu_2, \dots, \mu_n)'$, the vector of error terms as $e_{n \times 1} = (e_1, e_2, \dots, e_n)'$, and the matrix of regressors as $X_{n \times k} = (x_1, x_2, \dots, x_n)'$. If there are M model candidates $\{\mu^1, \mu^2, \dots, \mu^M\}$ to estimate the fitted values, the JMA estimator of μ takes the form $\mu = \sum_{j=1}^M \omega^j \mu^j$, where $\omega_{(1 \times 1)}^j$ is the weight assigned to the j th model using JMA approach and $\mu^j = (\mu_1^j, \mu_2^j, \dots, \mu_n^j)'$ is the vector of fitted values obtained from the j th model such that μ_i^j is computed with the i th observation deleted (known as leave-one-out cross-validation). For instance, if one uses OLS to estimate each model, then

$$\mu_i^j = x_i^{j'} (X_{(-i)}^{j'} X_{(-i)}^j)^{-1} X_{(-i)}^{j'} y_{(-i)}$$

where $X_{(-i)}^j$ and $y_{(-i)}$ represents matrices $X_{n \times k}^j$ and $y_{n \times 1}$ with the i th row deleted. Consequently, the Jackknife residual vector of the j th model is $e_{(n \times 1)}^j = y_{(n \times 1)} - \mu_{(n \times 1)}^j$ and the $n \times M$ matrix of residuals is $\Upsilon_{n \times M} = (e_{n \times 1}^1, e_{n \times 1}^2, \dots, e_{n \times 1}^M)$. The JMA choice of optimal weights minimizes

$$\begin{aligned} \min_{\omega_j} CV(\Omega) &= \frac{1}{n} \Omega' \Upsilon' \Upsilon \Omega \\ \text{s.t. } \sum_{j=1}^M \omega_j &= 1, \quad 0 \leq \omega_j \leq 1, \quad \forall j = 1, 2, \dots, M. \end{aligned} \tag{2.13}$$

where $\Omega_{(M \times 1)} = (\omega^1, \omega^2, \dots, \omega^M)'$ is the weight vector consisting of the weights assigned to the M models. Hansen & Racine (2009) theoretically prove JMA's asymptotic optimality and then illustrate their proof using a Monte Carlo simulation. They also provide an empirical example where they use a cross-section dataset consisting of observations on hourly wage and 30 explanatory variables. They empirically show that JMA has a better performance compared with its peers. To proceed with the empirical analysis, they consider 30 models from a pool of more than one billion potential models. This could cast doubt on the validity of the size of each variable's effect on predicted earnings. The orthogonalization of model space provides a way around this problem.

2.8 Model Space Orthogonalization

To implement the JMA estimator, the question of how best to select variables and combine them for all the models to be weighted within the cross-validation criterion is paramount. Here, I use the pioneering insight of Magnus et al. (2010) to reduce the model space from 2^k to k . This result is obtained by showing that the key results of Hansen & Racine (2009) do not change when I have an orthogonalization of the entire set of covariates.

Let X_m denote the matrix of covariates used to construct the OLS estimates for the m th model. The OLS estimator for the m th model is $\hat{\beta}^{(m)} = (X_m' X_m)^{-1} X_m' y$. The matrix $P_m = X_m (X_m' X_m)^{-1} X_m'$ will figure prominently in the following discussion. What I want

is to have X_m such that

$$Q'X'_mX_mQ = \Lambda, \quad (2.14)$$

Where Q is an orthogonal matrix and Λ is a diagonal matrix. Multiplying both sides of equation 2.14 by $\Lambda^{-1/2}$, I then have $\Lambda^{-1/2}Q'X'_mX_mQ\Lambda^{-1/2} = I$. Letting $\tilde{X}_m = X_mQ\Lambda^{-1/2}$ I have $\tilde{X}'_m\tilde{X}_m = I$. The m th Jackknife estimator is $\mu^m = (\mu_1^m, \mu_2^m, \dots, \mu_n^m)'$ and can be written as $\mu^m = P_m^{JMA}y$ such that

$$P_m^{JMA} = D_m(P_m - I) + I \quad (2.15)$$

where D_m is the diagonal matrix with the j th diagonal element equal to $(1 - h_{jj}^m)^{-1}$ and h_{jj}^m is the j th diagonal element of P_m . I note that all the results in Hansen & Racine (2009) still hold after orthogonalization because

$$\tilde{P}_m = \tilde{X}_m(\tilde{X}'_m\tilde{X}_m)^{-1}\tilde{X}'_m = \tilde{X}_m\tilde{X}'_m = X_mQ\Lambda^{-1}Q'X'_m = X_m(X'_mX_m)^{-1}X'_m = P_m.$$

This means the Jackknife projection matrix, P_m^{JMA} , remains same if I orthogonalize the entire model space. Also, it is trivial to recover $\beta^{(m)}$ from $\beta^{*(m)}$ (the coefficient vector from the orthogonalization) since

$$\tilde{X}_m\beta^{*(m)} = \tilde{P}_my = P_my = X_m\beta^{(m)}.$$

This yields $\beta^{*(m)} = \Lambda^{1/2}Q'\beta^{(m)}$ so once I have estimates of $\beta^{*(m)}$ I can recover estimates of the parameters of interest as $\beta^{(m)} = Q\Lambda^{-1/2}\beta^{*(m)}$.

The importance of the orthogonalization is that the residuals, $\hat{\varepsilon}_m = M_my = (I - P_m)y$ are identical across the k models that I estimate for the model averaging. Even more importantly, I can show that the models have the following structure $\hat{\varepsilon}'\hat{\varepsilon} = y'y - y'P_1y - \dots - y'P_my$. What the orthogonalization gives me is the ability to select a single variable and continue to add to the model space *linearly*. Also, given that the variables are orthogonal the order in which this is done is immaterial. Given that the cross-validation criteria can now be decomposed as the piecewise sums of the individual regressors, this makes sampling from the total set of models trivial. The simplest way to orthogonalize the vector space is to use the singular value decomposition or Gram-Schmidt orthogonalization.

Once the orthogonalization has been undertaken the JMA weights can be found by simply estimating the models by increasing the size of the model space by one each time.¹⁷ For example, suppose that there were 5 regressors which one wanted to average over. The total number of models to average over would be $2^5 = 32$. However, if I orthogonalize the regressors then I have only 5 models, $\{1, (1, 2), (1, 2, 3), (1, 2, 3, 4), (1, 2, 3, 4, 5)\}$, to average over. This may seem trivial for a small number of regressors, but it quickly becomes computationally relevant. With just 30 regressors, not uncommon in applied work, I already have more than a billion potential enumerations of the model space. Thus, JMA can be implemented across a large number of covariates without concern that enough models are visited, as in BMA investigations.

¹⁷Given the independence of the regressors the order in which this is done is irrelevant.

2.9 Data

My basis for assessing JMA performance is the Boston housing dataset, which is originally published in Harrison & Rubinfeld (1978) (HR hereafter) and is taken from the Census Tracts in the Boston Standard Metropolitan Area for the year 1970. This cross-section dataset consists of 506 observations along with 13 potential determinants of housing value in the Boston Metropolitan Area. The dependent variable (MV) is the corrected median value of owner-occupied homes in USD 1000's, and among the explanatory variables I have per capita crime rate ($CRIM$), the pupil-teacher ratio by town school district ($PTRATIO$), the index of accessibility to radial highways (RAD), the proportion of owner-occupied units built prior to 1940 (AGE), the proportion of non-retail business acres per town ($INDUS$), the proportion of African-Americans per town (B), and a variable reflecting the nitric oxides concentration (NOX). See Harrison & Rubinfeld (1978, Table IV) for more details. The authors use a hedonic housing price model to generate quantitative estimates of the willingness to pay for air quality improvements. I assume there exists uncertainty about which housing attributes should be a part of the final model. Orthogonalizing the model space, I employ JMA technique to uncover the true effect of each attribute on the value of homes. For comparison purposes, I also apply a Bayesian model averaging (BMA) analysis to determine the sets of explanatory variables, which are likely to make up the final model. By averaging across a large set of models, I can determine those variables which are relevant to the data generating process for a given set of priors used in the analysis. The next section presents the empirical results.

2.10 Results

For my analysis I construct BMA, and JMA estimators using HR dataset and compare the findings. For BMA implementation I use uniform distribution for model priors and set the hyperparameter on Zellner's g -prior for the regression coefficients to the benchmark prior suggested by Fernandez et al. (2001a). HR dataset is small enough that allows for full enumeration of model space and there is no need for using search algorithms to walk through the model space. Having said that, my BMA estimator visits $2^{13} = 8192$ models¹⁸. All this models include an intercept. My implementation of modified JMA starts with orthogonalizing the entire dataset, adding all variables linearly to the model, and computing the weights by minimizing the cross-validation criterion discussed in sections 2.7. Similarly, all the visited models in this setup include an intercept.

Table 2.5 reveals the BMA and JMA estimates for the Harrison & Rubinfeld (1978) study. Column (i) presents the Bayesian posterior inclusion probabilities (PIP) for the variables with PIPs above 50%. I have sorted the explanatory variables from the highest PIP to the lowest. Column (ii) presents the estimated posterior means and their standard errors, calculated by the BMA estimator. All the estimations appear with the expected signs and magnitude similar to those presented in Harrison & Rubinfeld (1978, Table VIII). Column

¹⁸Recall that 13 represents the total number of covariates in my dataset.

(iii) reports the estimated coefficients of interest and their corresponding standard errors using the original JMA estimator. The results of column (iii) are obtained by enumerating the entire model space, which means JMA has computed the optimal weights visiting 8192 models in my example. Column (iv), however, reports the corresponding model average estimates using the JMA estimator with the variable orthogonalization described in section 2.8. The number of models visited in this case is only 13 models while this number for columns (ii) and (iii) is $2^{13} = 8192$. As is apparent, orthogonalization drastically reduces the size of model space over which I conduct the Frequentist model averaging with no significant change in the outcomes. Comparing across the three set of estimates, I see that both the sign and magnitude of all three sets of estimates are almost identical.

2.11 Conclusion

Model uncertainty in statistical and econometric modeling rises when there are several explanatory variables that are likely a part of the data generating process while there is not a robust theoretical model evincing a clear link between the set of potential determinants and the dependent variable. Model averaging techniques are among the variety of methods to deal with model uncertainty. The main two categories of model averaging techniques are Bayesian model averaging and Frequentist model averaging. The former, which is perhaps the most popular model averaging technique as present, requires the choice of model priors and employs complex search algorithms such as Monte Carlo Markov Chain (MCMC) to search over the model space when the enumeration of the all possible combinations is computationally infeasible. The latter does not require a set of prior distributions for the parameters of the model but does require a set of appropriate weights assigned to the all candidate models. Researchers have recently proposed a wide array of criteria that can be used to obtain suitable weights for averaging the models within the Frequentist model averaging arena. For instance, Hansen (2007) proposes selecting the weights by minimizing a Mallows criterion and Hansen & Racine (2009) propose a Jackknife model averaging estimator (JMA) that selects the weights by minimizing a cross-validation criterion.

An issue that remains unsolved within the Frequentist model averaging is how to deal with a large model space. BMA methods use MCMC algorithms to reduce the model space but a similar algorithm within Frequentist framework, to my knowledge, has not been developed yet. This paper shows how orthogonalizing the entire dataset makes it possible to easily implement model averaging estimators in the presence of a large number of covariates and should prove useful in an array of economic contexts. I theoretically show that this innovation reduces the model space from 2^k to k where k is the number of covariates. In closing, an empirical application shows that the estimated coefficients and standard errors obtained from a JMA estimator in an orthogonalized environment are identical to those obtained from enumerating the entire model space.

This paper set out to replicate several important studies focused on model uncertainty in the modeling of economic growth. Current research on this influential area within growth empirics has almost solely relied on BMA techniques to date. Aside from my goal of repli-

Table 2.5: BMA, and JMA Estimation for Harrison & Rubinfeld (1978) Study.

	BMA	BMA	JMA	JMA
Method:	PIP	Post Mean/SD	Enumeration	Orthogonalization
	(i)	(ii)	(iii)	(iv)
Variable:				
Crime	1.00	-0.0119 (0.0012)	-0.0118 (0.0012)	-0.0117 (0.0014)
Lstat	1.00	-0.3705 (0.0228)	-0.3699 (0.0258)	-0.3644 (0.0321)
Dis	1.00	-0.1940 (0.0267)	-0.1902 (0.0288)	-0.1891 (0.0375)
Ptatio	1.00	-0.0314 (0.0047)	-0.0310 (0.0050)	-0.0305 (0.0056)
Nox	1.00	-0.0062 (0.0011)	-0.0063 (0.0011)	-0.0062 (0.0012)
Rm	0.99	0.0063 (0.0012)	0.0063 (0.0014)	0.0063 (0.0015)
Rad	0.99	0.0951 (0.0190)	0.0946 (0.0186)	0.0940 (0.0212)
Tax	0.98	-0.0004 (0.0001)	-0.0004 (0.0001)	-0.0004 (0.0001)
Blacks	0.96	0.3548 (0.1219)	0.3624 (0.1064)	0.3573 (0.1119)
Chas	0.69	0.0640 (0.0504)	0.0911 (0.0339)	0.0907 (0.0356)

Notes: This table presents the Bayesian model averaging (BMA), and Jackknife model averaging (JMA) techniques applied to the data used in Harrison & Rubinfeld (1978). The dependent variable (*MV*) is the corrected median value of owner-occupied homes in USD 1000's for the year 1970. Standard errors are in parenthesis. All models estimated contain an intercept. Column (i) reports the Bayesian posterior inclusion probabilities (PIPs), calculated by the BMA estimator. Column (ii) reports the Bayesian posterior means and their standard errors. Column (iii) reports the estimation of model coefficients and standard deviations using the JMA approach discussed in Hansen & Racine (2009). Column (iv) reports the corresponding model averaging estimates obtained from the JMA estimator using the orthogonalization of model space described in Magnus et al. (2010).

cation I also compared the BMA results to similar estimates obtained using alternative Bayesian methods (WALS) and recently introduced FMA techniques (MMA). These newer methods may be appealing to some because they allow for full enumeration of the model space which avoids having to invoke MCMC search algorithms to sample from the model space. Additionally, the FMA methods do not require the specification of prior distributions to make the averaging estimator operational. For my FMA estimator I employed

the Mallows's criterion to select the weights as it has been shown to deliver asymptotically optimal weights. Using freely available software to implement the Bayesian methods I was able to exactly reproduce the results of all three studies. Additionally, I found that the sign and magnitude of both the WALS and the MMA coefficients were roughly identical to those obtained via BMA.

I am encouraged by the similarity of these three methods as well as my ability to replicate the results of these computationally intensive estimation methods. However, several important caveats are noted. The primary focus of my replication/comparison was based on the weighted coefficient estimates and their standard errors. However, traditional BMA analyses also focus attention on the PIP of a variable as well as measuring jointness between pairs of covariates. In the current setup, neither WALS or MMA provide this information. This suggests an extant relative strength of BMA over its competitors. Future work to derive similar metrics from WALS and MMA are required before having the ability to favor one method over another categorically.

Beyond my comparison and replication of three important studies into the determinants of economic growth, my implementation of the MMA estimator required use of orthogonalization of the covariate space to fully enumerate the model space, making the analyses computationally feasible. This orthogonalization is the same insight that makes WALS appealing in an applied context. Orthogonalization makes it possible to easily implement model averaging estimators in the presence of a large number of covariates and should prove useful in an array of economic contexts.

Chapter 3

Weighting to be Happy: A Study of Determinants of Happiness

3.1 Abstract

This paper explores a wide range of cross-country determinants of happiness and investigates the issue of model uncertainty in happiness and well-being regressions using Bayesian Model Averaging (BMA). BMA, which is firmly established in statistical theory, enables me to consider a large number of potential explanatory variables and sort out which of these variable can effectively explain what makes people happy. Posterior inclusion probabilities along with coefficient estimates reveal that income, being employed, general health, being socially active, having a partner, not being discriminated against, and being involved in volunteering activities are among the key determinants of happiness. The degree of dependence (jointness) among explanatory variables is also examined. Looking at jointness scores uncovers unknown forms of dependence among regressors and determines the regressors that are substitutes and capture similar underlying effects and the regressors that are complements and represent mutually reinforcing effects. The robustness of results is tested over different waves of the dataset.

3.2 Introduction

Improving the welfare of the nation is definitely the most challenging goal that politicians seek to achieve. Not surprisingly, it is not easy to find ready-made definitions of social welfare that can be used to elaborate and properly evaluate policy interventions. Starting from the famous Easterlin's (2001) paradox that questions the expected positive correlation between income per capita and subjective well-being, the classic economic dichotomy between GDP and citizens welfare has been (and continues being) the object of a critical debate. As an established result (see Frey & Stutzer 2002, Ferrer-i-Carbonell & Frijters 2004, Bruni & Porta 2005, Dolan, Peasgood & White 2008, for surveys), "money is not enough to make people happy" and in addition to economic factors such as income, low inflation rate (Di Tella, MacCulloch & Oswald 2001), and employment conditions (Di Tella et al. 2001, Clark & Oswald 1994), there are noneconomic variables that are found to play a crucial role in stimulating life satisfaction: being educated (Easterlin 2001, Blanchflower & Oswald 2004), enjoying good health conditions (Frey & Stutzer 2001), being female (Easterlin 2001, Shmotkin 1990), not being discriminated against (Davis 1984, Hughes & Thomas 1998), being native rather than immigrant (Frey & Stutzer 2001), being active in the social life, volunteering, and having frequent contacts with family, friends and neighbors (Bjørnskov 2003, Helliwell & Putnam 2004), being married and having children (Diener, Gohm, Suh & Oishi 2000, Stack & Eshleman 1998, Kohler, Behrman & Skytthe 2005, Waite 1995), and praying and attending religious communities (Ellison 1991) are considered as important non economic variables which positively influence the level of self-reported happiness..

Investigating into the determinants of life satisfaction does not simply represent an academic field of study. Rather, in elaborating their policy interventions, politicians start exhibiting an increasing interests in the empirical studies on subjective well-being elaborated by economists, sociologists and psychologists. For instance, Nicolas Sarkozy, the former French president, established in 2009 the Commission on the Measurement of Economic Performances and Social Progress (the so called Stiglitzs commission) with the intent to define a multidimensional welfare index that, in addition to income and financial wealth, included the other main non economic determinants of life satisfaction.¹

Empirical studies on subjective well-being seek to identify the relation between life satisfaction as self-reported by respondents participating in general surveys and a set of subjective characteristics. Typically, the choice of which determinants should be included in the set of independent variables is quite discretionary as it depends on the research question of the researcher as well as the objective structure of the questionnaire. Despite of its importance, defining a criterion to both rank the determinants in terms of their explanatory power and choose which variables effectively enter in the domain of life satisfaction still remains an open question. Identifying the relation between life satisfaction (or other proxy of utility

¹In its 2009 report, available at www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf, the commission indicates that "Steps should be taken to improve measures of peoples health, education, personal activities and environmental conditions. In particular, substantial effort should be devoted to developing and implementing robust, reliable measures of social connections, political voice, and insecurity that can be shown to predict life satisfaction" (page 15).

such as happiness) and economic, demographic, social and physical characteristics of individuals is the main goal of studies on happiness. In particular, in order to collect data on individuals well-being, researchers use general surveys asking individuals to self-report their life satisfaction on a scale which goes from a minimum value to indicate very low satisfaction to a maximum value which expresses very high satisfaction.

There is no doubt about the importance of these studies. Starting from the famous Easterlin's (2001) paradox which states the existence of a reversed U-shaped relation between happiness and income in the US between 1946-1996, social scientists have devoted a large quantity of research using self-reported data to identify the main determinants of life satisfaction. In this study I seek to employ Bayesian model averaging analysis to uncover the main determinants of happiness. The reminder of the paper is organized as follows. Section 3.3 reviews the implementation of BMA framework. Section 3.4 describes the nature of dataset I am using in this project while section 3.5 presents the main results of the paper by applying BMA method to the potential determinants of life satisfaction. Section 3.6 provides the degree of jointness between the regressors using common jointness scores and section 3.7 reports the BMA results obtained from individual countries and highlights the differences between the main factors of happiness in different European countries. Section 3.8 discusses the results and concludes.

3.3 Model Uncertainty and Bayesian Model Averaging Analysis

The majority of empirical studies on socio-economic phenomena such as economic growth or happiness face the problem of model uncertainty. This issue emerges because the relevant theories do not tell me what the true model is and what variables belong in the model. Over the last two decades, statisticians and econometricians have attempted to address the model uncertainty issue by developing statistical methods that allow for systematic search over all possible models and identify those that are more likely to be the true data generating processes. These methods, called model averaging techniques, appropriately incorporate model uncertainty into empirical analyses and are capable of dealing with the multiplicity of possible regressors. Model averaging methods treat models and related parameters as unobservable and estimate the distribution of these unknown parameters by allowing for all possible models to be considered. In contrast to model selection methods, model averaging procedures provide better predictive performance than any single selected model in the presence of model uncertainty (see Raftery et al. 1997). Bayesian Model Averaging (BMA hereafter) and Frequentist Model Averaging are the most promising approaches to model averaging. This section provides a brief intuition behind the BMA methodology.

Suppose we have k possible explanatory variables, $x_1, x_2, \dots, x_k$. Therefore, there are 2^k possible models, $M = \{M_1, M_2, M_3, \dots, M_{2^k}\}$, that all can potentially explain the data. Considering the j th model, M_j , a generic linear regression model under M_j is

$$y = X^j \beta^j + \epsilon \quad (3.1)$$

where y is the vector of observations on the variable of interest, X^j is the matrix of all explanatory variables considered for the j th model, β^j is the vector of parameters of model M_j , and ϵ is the standard error term. The probability distribution of β^j given j th model, M_j and data, Δ , is

$$p(\beta^j|\Delta, M_j) = p(\Delta|\beta^j, M_j) \frac{p(\beta^j|M_j)}{p(\Delta|M_j)} \quad (3.2)$$

This probability distribution can be used to measure the probability that M_j actually generated the data. Using Bayes' rule

$$p(M_j|\Delta) = p(\Delta|M_j) \frac{p(M_j)}{p(\Delta)} \quad (3.3)$$

where $p(M_j)$ is the prior model probability, and $p(\Delta|M_j)$ is the marginal or integrated likelihood. Practical use of this equation requires few manipulations. As is apparent, $p(M_j)$ does not depend on the data and shows my prior belief that M_j is the correct model before seeing the data. To calculate the marginal likelihood, one can integrate both sides of equation (3.2) with respect to β^j yielding

$$\begin{aligned} \int p(\beta^j|\Delta, M_j) d\beta^j &= \int p(\Delta|\beta^j, M_j) \frac{p(\beta^j|M_j)}{p(\Delta|M_j)} d\beta^j \\ \Rightarrow 1 &= \frac{1}{p(\Delta|M_j)} \int p(\Delta|\beta^j, M_j) p(\beta^j|M_j) d\beta^j \\ \Rightarrow p(\Delta|M_j) &= \int p(\Delta|\beta^j, M_j) p(\beta^j|M_j) d\beta^j \end{aligned} \quad (3.4)$$

where $p(\beta^j|M_j)$ is prior probability distribution assigned to the parameters of model M_j . Whether or not the data favors M_i over M_j depends on the Bayes factor, $\frac{p(\Delta|M_i)}{p(\Delta|M_j)}$. When this ratio is greater than 1, the data supports M_i over M_j . Finally, $p(\Delta)$, the last piece of equation (3.3), can be computed using the law of total probability as follows

$$p(\Delta) = \sum_{i=1}^{2^k} p(\Delta|M_i) p(M_i) \quad (3.5)$$

The logic of BMA, then, suggests that the posterior density of $\beta = (\beta_1, \beta_2, \dots, \beta_k)'$ in full model $y = X\beta + \epsilon = \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \epsilon$ is

$$p(\beta|\Delta) = \sum_{i=1}^{2^k} p(\beta|\Delta, M_i) p(M_i|\Delta) \quad (3.6)$$

Taking expectations across (3.6) gives the posterior mean for parameter β as

$$E(\beta|\Delta) = \sum_{i=1}^{2^k} E(\beta|\Delta, M_i) p(M_i|\Delta) \quad (3.7)$$

Following Leamer (1978), the posterior variance of parameter of interest would be

$$Var(\beta|\Delta) = \sum_{i=1}^{2^k} Var(\beta|\Delta, M_i)p(M_i|\Delta) + \sum_{i=1}^{2^k} (E(\beta|\Delta, M_i) - E(\beta|\Delta))^2 p(M_i|\Delta) \quad (3.8)$$

A useful feature of BMA is that it allows the estimation of the posterior probability that a particular explanatory variable, let say x_h , belongs to the true model. This probability is the sum of the posterior probabilities for all the models that include x_h . Mathematically,

$$p(x_h|\Delta) = \sum_{i: x_h \in M_i} p(M_i|\Delta) \quad (3.9)$$

If posterior inclusion probability (hereafter PIP) obtained from (3.9) is high for variable x_h , this variable is considered as a robust determinant of variable of interest. While BMA is an intuitively attractive solution to be used to overcome the problem of model uncertainty in empirical research, its implementation presents a number of difficulties and challenges. For instance, the size of model space increases exponentially as k , the number of explanatory variables, goes up. This can render the summations computationally infeasible or expensive in the previous equations. To deal with this issue, various algorithms are proposed in the literature. The Markov Chain Monte Carlo (MCMC) algorithm, to my knowledge, is the most widely used technique to search over the model space. I employ the same methodology in this work to search in the pool of potential determinants of happiness.

Another controversial aspect of implementing BMA is the specification of the prior model probabilities $p(M_j)$ as well as the parameters' prior probability distributions, $p(\beta^j|M_j)$. The set of prior probabilities must be informative with respect to the likelihood, meaning priors should be relatively high where the likelihoods are large otherwise the choice of priors will have a substantial effect on the posteriors.

Several options exist for the specification of model priors $p(M_j)$. Initially proposed by Fernandez et al. (2001a), a common option for the specification of the model priors is a uniform distribution over the model space, implying that the researcher has no preference for a specific model. Accordingly, the prior probability assigned to each model will be

$$p(M_1) = p(M_2) = \dots = p(M_{2^k}) = 1/2^k \quad (3.10)$$

Throughout this research, I will implement BMA, assuming uniform distribution for the model priors. Other common model priors proposed in the literature are those that penalize high number of regressors in the model (complexity). Sala-i-Martin et al. (2004), for example, employ a prior model distribution that reflects the researcher's prior beliefs on the size of the model.

Following the Bayesian Model Averaging literature, I assume a normal distribution as the choice of prior structures for the coefficients. More specifically, I assume

$$p(\beta^j|M_j) \sim N(0, \sigma^2\Omega) \quad (3.11)$$

where prior variance matrix Ω is assumed to be proportional to the sample covariance,

$$\Omega = (gX^{j'}X^j)^{-1} \quad (3.12)$$

The choice of the factor of proportionality in this expression, g , which was first suggested by Zellner (1986), is crucial in BMA analyses. For example, Fernandez et al. (2001a) recommend setting $g = n$, or $g = \max(n, k^2)$ where n is the sample size and k is the number of regressors, Foster & George (1994) introduce a risk inflation criterion and set $g = k^2$, and George & Foster (2000) and Liang et al. (2008) estimate a local empirical Bayes g -parameter.

3.4 Data

I use a survey dataset that includes more than 185,000 observations on different potential determinants of happiness across 28 European countries. There are 4 waves (2002, 2004, 2006, 2008) and the collecting procedure lasts one year for each wave. The dataset has a cross-section nature, meaning that individuals are not followed over time and new subjects fill the questionnaire in each wave. My dataset consists of information on

1. Demographic variables such as
 - (a) Gender
 - (b) Age
 - (c) Whether or not the respondent is born in the country area
 - (d) Whether or not the respondent belongs to a minority ethnic group
 - (e) Whether or not the respondent belongs to a group that is discriminated against
2. Income and economic status including
 - (a) Total income
 - (b) How easily the respondent is able to borrow money
 - (c) Whether or not the respondent has been a victim of burglary in the last 5 years
 - (d) How the respondent feels about the household's income
3. Civil status and family background, for instance
 - (a) Number of household members
 - (b) Whether or not the respondent is married, divorced, separated, widowed, or has a partner
 - (c) If the respondent lives with children in the house
4. Employment status such as

- (a) Whether or not the respondent is unemployed, retired, has a paid work, works from house, has an education job, is self-employed, has a family-owned business, and enjoys an unlimited job contract
- (b) Total number of hours worked per week
- (c) If the respondent supervises other employees
- 5. Health including
 - (a) How the respondent feels about her general health
 - (b) If the respondent has limitations with her daily activities
- 6. Political background, for example
 - (a) If the respondent is politically left oriented or right oriented
 - (b) Respondent's placement on left and right scale in politics
 - (c) If the respondent has participated in the last election
 - (d) How interested the respondent in politics is
- 7. Religious background such as
 - (a) How religious the respondent is
 - (b) If the respondent belongs to a particular religion or denomination
 - (c) How often the respondent prays
 - (d) If the respondent is Jewish, Muslim, Roman-Catholic, Protestant, Easternorthodox, etc.
- 8. Education background including
 - (a) Number of years of schooling
 - (b) Education level
 - (c) If the respondent has improved her skills in the last 12 months or attended a useful lecture
- 9. Interpersonal relationships such as
 - (a) How often the respondent socializes with here friends and colleagues
 - (b) If the respondent has somebody to talk to about intimate and personal matters
- 10. Social activities such as
 - (a) Whether or not the respondent voluntarily has worked in the last 12 months in an organization or association
 - (b) Whether or not the respondent has worked in a political party in the last 12 months

There are 4 waves available in the dataset and the dependent variable is the overall level of respondent's life satisfaction. Table 3.1 provides summary descriptive statistics for the data.

3.5 BMA results

Now, I employ the BMA estimator and apply it to different waves of my dataset to compute the PIPs, coefficients, and standard errors. As discussed, implementing the BMA estimator requires specifying the set of model priors and a search algorithm. I use the uniform distribution for model priors and a MCMC sampler to search over the model space to construct my BMA estimator. I also use the *g*-prior framework, introduced in Fernandez et al. (2001a), as the benchmark for prior assumptions of the model parameters. Table 3.2 presents the posterior inclusion probabilities (PIP), posterior means (Post Mean), and posterior standard deviations (Post SD) obtained by applying BMA to the wave 1. The explanatory variables are sorted from the highest PIP to the lowest. Columns (i) & (iv) report the PIPs, columns (ii) & (v) show the estimated posterior means, and columns (iii) & (vi) present the estimated posterior standard errors.

As is present, all the determinants of happiness with a PIP higher than 50% are statistically significant and have the expected sign. For example, whether or not the respondent belongs to a discriminated group in the country (Discriminated) appears with a posterior inclusion probability of 100% and a statistically significant and negative coefficient. This implies that people who are discriminated against feel less happy controlling for other potential factors. Income variable in my dataset is a point-scale variable ranging from 1 to 12. The higher the value of this scale is, the more income is earned by the respondent. My analysis shows that how much money a person is making is definitely a key determinant of happiness. The coefficient on income is positive and statistically significant. People who believe that money does not bring happiness may not like this finding but looking at the results shows that the effect of income on happiness (its magnitude) is much less than almost all other important determinants of happiness that have a PIP greater than 50%. Being discriminated against, for instance, has ten times larger impact on how happy we are than the impact of a one-point change in how much money we make. The left column of table 3.2 demonstrates that having a job (Unemployed), being separated or having a partner, affiliated with a religion, being healthy, working hours, having high level of education, being socially involved, and having someone to talk to are other important factors that determine how happy we are. Looking at the bottom numbers of right column of table 3.2 reveals that being married, divorced, or widowed, surprisingly, have little effect on happiness.

As a robustness check, I run the same analysis using other available waves in my dataset, wave 2, wave 3, waves 1 through 3 combined, and waves 1 through 4 combined. The results are shown in tables 3.3, 3.4, 3.6, and 3.7, respectively. My findings show that the main determinants of happiness shown in table 3.2 are robust and remain statistically significant with the expected signs.

Tables 3.8, 3.9, 3.10, 3.11, 3.12, and 3.13 show the estimated PIPs as well as the post mean & standard errors dropping the subjective variables.

3.6 Jointness of Happiness Determinants

Capturing the dependence between explanatory variables in the posterior distribution while implementing a Bayesian analysis is crucial. Taking such a dependence into account reveals the sensitivity of posterior distributions of parameters to dependence across regressors. For instance, if two specific variables are complementary over the model space, I expect relatively higher weights for the models that include both variables together than models where only one variable is present. To introduce my measure of dependence/jointness, I follow Doppelhofer & Weeks (2009) and Ley & Steel's (2009b) methodology.

Equation (3.9) provides a measure for the marginal posterior inclusion probability of a given regressor (x_h in this case). According to this equation, the posterior probability of x_h can be calculated by summing the posterior weights of models that include that explanatory variable. Although the marginal probability measure of variable importance has interesting policy implications and prediction, it hardly provides information on the nature of dependencies among regressors. To address this issue, I first introduce the posterior joint inclusion probability of variables x_h and x_l as

$$p(x_h \cap x_l | \Delta) = \sum_{i: x_h \& x_l \in M_i} p(M_i | \Delta) \quad (3.13)$$

Similar to (3.9), this equation adds up posterior model probabilities for all the models in which the two variables are present. Consequently, a natural measure of dependence between variables x_h and x_l is given by²

$$J_{hl} = \ln \left[\frac{p(x_h \cap x_l | \Delta) \cdot p(\bar{x}_h \cap \bar{x}_l | \Delta)}{p(\bar{x}_h \cap x_l | \Delta) \cdot p(x_h \cap \bar{x}_l | \Delta)} \right] \quad (3.14)$$

Positive value for this jointness statistic indicates that variables x_h and x_l are *complements*, reflecting that posterior probabilities of models jointly including & excluding both variables are larger than posterior probabilities of models including only one. A negative value for the statistic, on the other hand, indicates that the variables are capturing a similar underlying effect and are called *substitutes*. The degree of dependence or the significance of jointness can be defined as

$$\text{Dependence degree : } \begin{cases} -\infty < J_{hl} < -2 & \text{Strong substitute} \\ -2 < J_{hl} < -1 & \text{Significant substitute} \\ -1 < J_{hl} < +1 & \text{Not significantly related} \\ +1 < J_{hl} < +2 & \text{Significant complements} \\ +2 < J_{hl} < +\infty & \text{Strong complements} \end{cases}$$

As is present in (3.14), the behavior of this jointness statistics depends on whether or not the posterior inclusion probability of one of the explanatory variables is zero or extremely small, a case that commonly occurs in practice. Ley & Steel (2009b) modify the jointness

²For more details, see Doppelhofer & Weeks (2009) and Whittaker (1990, Ch. 2)

statistic and consider the following measures of jointness

$$J_{hl}^o = \ln \left[\frac{p(x_h \cap x_l | \Delta)}{p(x_h \cup x_l | \Delta)} \right] \quad (3.15)$$

and

$$J_{hl}^* = \ln \left[\frac{p(x_h \cap x_l | \Delta)}{p(\bar{x}_h \cap x_l | \Delta) + p(x_h \cap \bar{x}_l | \Delta)} \right] \quad (3.16)$$

They show these two statistics, that have formal statistical and intuitive meaning in terms of jointness, are more likely to be well-behaved when the posterior inclusion probability of one of the regressors is either zero or very close to zero. My choice of jointness statistic in this study is (3.16). Table 3.14 and 3.15 present results of applying jointness, defined in equation (3.16), to the determinants of happiness across wave 1. Given the degree of dependence defined in section 3.6, I find evidence of both positive and negative jointness among happiness regressors. Recall that positive jointness score indicates that two examined variables are complement while negative jointness score indicates that two variables are substitutes.³ For example, *Unemployed* and *Discriminated* are strong complements to each other, implying that the size and significance of the coefficient of *Discriminated* is strengthened conditional on controlling for *Unemployed*. On the other hand, *Age* and *Widowed*, for instance, exhibit strong negative jointness meaning that one loses its significance once the other one is included as regressor in the model.

To examine the robustness of results of jointness scores, I apply the jointness measure to other waves of dataset. Results for waves 2, 3, 4, 1-3, and 1-4 are presented in tables 3.16, 3.17, 3.18, 3.19, 3.20, 3.21, 3.22, 3.23, 3.24, and 3.25. My findings show robust jointness results across different waves of my dataset.

³'PC' in the jointness tables stands for Perfect Complement and 'PS' stands for Perfect Substitute.

Table 3.1: Data Descriptive Statistics

Variable	Nbr.Observations	Mean	St.Dev	Median	Min	Max	Range
Life Satisfaction	183808.00	6.80	2.36	7.00	0.00	10.00	10.00
Gender	184745.00	0.46	0.50	0.00	0.00	1.00	1.00
Age	183867.00	47.06	18.46	46.00	13.00	123.00	110.00
Country born	184699.00	0.91	0.28	1.00	0.00	1.00	1.00
Minority ethnic group	181462.00	0.05	0.22	0.00	0.00	1.00	1.00
Discriminated	181620.00	0.06	0.25	0.00	0.00	1.00	1.00
Income	129344.00	5.90	2.68	6.00	1.00	12.00	11.00
Borrow money	174955.00	2.74	1.25	3.00	1.00	5.00	4.00
Victim of burglary	184324.00	0.19	0.39	0.00	0.00	1.00	1.00
Income feeling	179362.00	2.07	0.88	2.00	1.00	4.00	3.00
Number of household members	184784.00	2.82	1.49	3.00	1.00	22.00	21.00
Partner	183418.00	0.60	0.49	1.00	0.00	1.00	1.00
Widowed	182309.00	0.10	0.30	0.00	0.00	1.00	1.00
Divorced	182309.00	0.07	0.26	0.00	0.00	1.00	1.00
Separated	182309.00	0.01	0.12	0.00	0.00	1.00	1.00
Married	182309.00	0.53	0.50	1.00	0.00	1.00	1.00
Living with children	184089.00	0.39	0.49	0.00	0.00	1.00	1.00
Unemployed	184320.00	0.06	0.24	0.00	0.00	1.00	1.00
Retired	184320.00	0.25	0.43	0.00	0.00	1.00	1.00
Paid work	184320.00	0.52	0.50	1.00	0.00	1.00	1.00
Community or military job	184320.00	0.00	0.06	0.00	0.00	1.00	1.00
Work from house	184320.00	0.21	0.41	0.00	0.00	1.00	1.00
Doing other things	184320.00	0.02	0.14	0.00	0.00	1.00	1.00
Education job	184320.00	0.10	0.30	0.00	0.00	1.00	1.00
Self-employed	159038.00	0.12	0.32	0.00	0.00	1.00	1.00
Family-owned business	159038.00	0.01	0.12	0.00	0.00	1.00	1.00
Unlimited job contract	176280.00	0.59	0.49	1.00	0.00	1.00	1.00
Work hours	150016.00	40.51	14.37	40.00	0.00	555.00	555.00
Supervising other employees	163254.00	0.29	0.46	0.00	0.00	1.00	1.00
General health	184718.00	2.24	0.93	2.00	1.00	5.00	4.00
Daily activities limitation	183927.00	0.25	0.43	0.00	0.00	1.00	1.00
Left political orientation	157758.00	0.32	0.47	0.00	0.00	1.00	1.00
Right political orientation	157758.00	0.35	0.48	0.00	0.00	1.00	1.00
Politics scale	157758.00	5.11	2.21	5.00	0.00	10.00	10.00
Voted last election	183008.00	0.73	0.44	1.00	0.00	1.00	1.00
Interested in politics	184429.00	2.62	0.90	3.00	1.00	4.00	3.00
Religious	183418.00	4.84	2.97	5.00	0.00	10.00	10.00
Belong to a religion	181419.00	0.64	0.48	1.00	0.00	1.00	1.00
Pray	181623.00	4.50	2.47	6.00	1.00	7.00	6.00
Roman Catholic	121820.00	0.45	0.50	0.00	0.00	1.00	1.00
Protestant	121820.00	0.19	0.39	0.00	0.00	1.00	1.00
Easternorthodox	121820.00	0.13	0.34	0.00	0.00	1.00	1.00
Other Christian	121820.00	0.02	0.14	0.00	0.00	1.00	1.00
Jewish	121820.00	0.03	0.17	0.00	0.00	1.00	1.00
Muslim	121820.00	0.06	0.23	0.00	0.00	1.00	1.00
Eastern religion	121820.00	0.00	0.06	0.00	0.00	1.00	1.00
Other non-Christian	121820.00	0.00	0.06	0.00	0.00	1.00	1.00
Years of study	182739.00	11.85	4.12	12.00	0.00	56.00	56.00
Years of study (squared)	182739.00	157.48	102.65	144.00	0.00	3136.00	3136.00
Education level (1)	184333.00	0.15	0.36	0.00	0.00	1.00	1.00
Education level (2)	184333.00	0.20	0.40	0.00	0.00	1.00	1.00
Education level (3)	184333.00	0.37	0.48	0.00	0.00	1.00	1.00
Education level (4)	184333.00	0.02	0.15	0.00	0.00	1.00	1.00
Education level (5)	184333.00	0.25	0.43	0.00	0.00	1.00	1.00
Education level (other)	184333.00	0.00	0.04	0.00	0.00	1.00	1.00
Improved skills	183791.00	0.30	0.46	0.00	0.00	1.00	1.00
Socializing	184146.00	4.93	1.60	5.00	1.00	7.00	6.00
Intimate	182957.00	0.90	0.31	1.00	0.00	1.00	1.00
Volunteer work	184040.00	0.13	0.34	0.00	0.00	1.00	1.00
Worked in a political party	184231.00	0.04	0.20	0.00	0.00	1.00	1.00
Member of a political party	184234.00	0.05	0.22	0.00	0.00	1.00	1.00
Social activities	180674.00	2.71	0.95	3.00	1.00	5.00	4.00
Attending religious services	183862.00	5.35	1.56	6.00	1.00	7.00	6.00
Trade union member	183593.00	0.19	0.40	0.00	0.00	1.00	1.00

Table 3.2: Bayesian Model Averaging Analysis: Wave 1

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Discriminated	1.0000	-0.4876	0.0350	Interested in politics	0.4520	0.0155	0.0185
Income	1.0000	0.0408	0.0041	Improved skills	0.3629	0.0239	0.0342
Borrow money	1.0000	0.1139	0.0073	Gender	0.3303	-0.0186	0.0285
Victim of burglary	1.0000	-0.1155	0.0208	Self-employed	0.2887	0.0238	0.0402
Income feeling	1.0000	-0.6570	0.0127	Eastern religion	0.2340	0.0815	0.1587
Partner	1.0000	0.2557	0.0203	Divorced	0.0963	-0.0083	0.0275
Separated	1.0000	-0.4054	0.0668	Member of a political party	0.0800	0.0070	0.0259
Unemployed	1.0000	-0.6010	0.0468	Work from house	0.0524	0.0026	0.0122
Paid work	1.0000	-0.2263	0.0334	Roman Catholic	0.0455	0.0037	0.0188
General health	1.0000	-0.5418	0.0106	Education level (1)	0.0442	0.0035	0.0186
Politics scale	1.0000	0.0617	0.0042	Supervising other employees	0.0430	-0.0017	0.0091
Religious	1.0000	0.0884	0.0043	Worked in a political party	0.0316	-0.0028	0.0181
Protestant	1.0000	0.4303	0.0259	Unlimited job contract	0.0298	-0.0011	0.0073
Easternorthodox	1.0000	-0.4539	0.0401	Living with children	0.0274	-0.0010	0.0069
Jewish	1.0000	0.4373	0.0602	Education level (4)	0.0272	-0.0028	0.0186
Socializing	1.0000	0.1009	0.0061	Left political orientation	0.0172	-0.0008	0.0069
Intimate	1.0000	0.3711	0.0303	Daily activities limitation	0.0139	-0.0005	0.0052
Social activities	1.0000	0.1117	0.0098	Attending religious services	0.0129	0.0001	0.0014
Trade union member	1.0000	0.1214	0.0212	Voted last election	0.0124	0.0004	0.0047
Pray	1.0000	0.0248	0.0045	Doing other things	0.0115	-0.0011	0.0126
Retired	0.9998	0.1918	0.0391	Number of household members	0.0096	-0.0001	0.0013
Work hours	0.9997	-0.0033	0.0007	Right political orientation	0.0084	-0.0002	0.0037
Volunteer work	0.9990	0.1124	0.0233	Education level (3)	0.0079	0.0002	0.0059
Education level (5)	0.9843	-0.1148	0.0302	Education level (2)	0.0079	0.0003	0.0070
Other Christian	0.9128	0.2118	0.0876	Widowed	0.0073	0.0002	0.0040
Muslim	0.8905	-0.1683	0.0745	Belong to a religion	0.0055	-0.0002	0.0060
Family-owned business	0.8601	0.2328	0.1148	Other non-Christian	0.0055	0.0002	0.0091
Country born	0.7142	-0.0775	0.0554	Married	0.0051	0.0000	0.0025
Age	0.7105	0.0022	0.0016	Education level (other)	0.0045	0.0001	0.0108
Education job	0.7081	0.1181	0.0854	Community or military job	0.0037	0.0003	0.0111
Years of study (squared)	0.5790	-0.0008	0.0007	Minority ethnic group	0.0037	-0.0001	0.0029
Years of study	0.5598	0.0187	0.0179				

Table 3.3: Bayesian Model Averaging Analysis: Wave 2

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Discriminated	1.0000	-0.4876	0.0350	Interested in politics	0.4584	0.0158	0.0187
Income	1.0000	0.0408	0.0041	Improved skills	0.3787	0.0250	0.0346
Borrow money	1.0000	0.1139	0.0073	Gender	0.3424	-0.0193	0.0288
Victim of burglary	1.0000	-0.1155	0.0208	Self-employed	0.2931	0.0241	0.0403
Income feeling	1.0000	-0.6569	0.0127	Eastern religion	0.2420	0.0847	0.1615
Partner	1.0000	0.2557	0.0203	Divorced	0.1038	-0.0090	0.0286
Separated	1.0000	-0.4056	0.0668	Member of a political party	0.0774	0.0067	0.0255
Unemployed	1.0000	-0.6006	0.0468	Work from house	0.0542	0.0027	0.0123
Paid work	1.0000	-0.2258	0.0333	Roman Catholic	0.0517	0.0043	0.0201
General health	1.0000	-0.5418	0.0106	Education level (1)	0.0501	0.0036	0.0177
Politics scale	1.0000	0.0616	0.0042	Supervising other employees	0.0442	-0.0018	0.0094
Religious	1.0000	0.0884	0.0043	Unlimited job contract	0.0321	-0.0012	0.0075
Pray	1.0000	0.0248	0.0045	Worked in a political party	0.0293	-0.0027	0.0176
Protestant	1.0000	0.4306	0.0266	Education level (4)	0.0280	-0.0028	0.0188
Easternorthodox	1.0000	-0.4529	0.0407	Living with children	0.0217	-0.0008	0.0061
Jewish	1.0000	0.4373	0.0605	Left political orientation	0.0188	-0.0009	0.0072
Socializing	1.0000	0.1010	0.0061	Daily activities limitation	0.0164	-0.0006	0.0056
Intimate	1.0000	0.3711	0.0303	Doing other things	0.0144	-0.0014	0.0141
Social activities	1.0000	0.1115	0.0098	Attending religious services	0.0142	0.0002	0.0015
Trade union member	1.0000	0.1212	0.0213	Voted last election	0.0140	0.0005	0.0052
Retired	0.9998	0.1915	0.0390	Number of household members	0.0111	-0.0001	0.0013
Work hours	0.9996	-0.0032	0.0007	Right political orientation	0.0077	-0.0002	0.0035
Volunteer work	0.9992	0.1124	0.0232	Education level (3)	0.0065	-0.0000	0.0033
Education level (5)	0.9862	-0.1137	0.0297	Widowed	0.0063	0.0001	0.0035
Other Christian	0.9124	0.2126	0.0882	Minority ethnic group	0.0056	-0.0001	0.0036
Muslim	0.8695	-0.1636	0.0777	Community or military job	0.0054	0.0004	0.0135
Family-owned business	0.8555	0.2313	0.1157	Belong to a religion	0.0050	-0.0001	0.0047
Age	0.7253	0.0023	0.0016	Education level (2)	0.0043	0.0001	0.0032
Education job	0.7221	0.1210	0.0848	Education level (other)	0.0040	0.0001	0.0103
Country born	0.7115	-0.0773	0.0555	Other non-Christian	0.0039	0.0002	0.0080
Years of study (squared)	0.6360	-0.0009	0.0007	Married	0.0039	0.0000	0.0021
Years of study	0.6159	0.0207	0.0177				

Table 3.4: Bayesian Model Averaging Analysis: Wave 3

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Discriminated	1.0000	-0.4876	0.0350	Interested in politics	0.4444	0.0152	0.0184
Income	1.0000	0.0408	0.0041	Improved skills	0.3557	0.0234	0.0339
Borrow money	1.0000	0.1139	0.0073	Gender	0.3478	-0.0196	0.0289
Victim of burglary	1.0000	-0.1156	0.0208	Self-employed	0.2816	0.0232	0.0398
Income feeling	1.0000	-0.6571	0.0127	Eastern religion	0.2360	0.0824	0.1596
Partner	1.0000	0.2556	0.0203	Divorced	0.0980	-0.0084	0.0277
Separated	1.0000	-0.4053	0.0668	Member of a political party	0.0837	0.0073	0.0264
Unemployed	1.0000	-0.6010	0.0468	Work from house	0.0551	0.0027	0.0124
Paid work	1.0000	-0.2261	0.0333	Roman Catholic	0.0496	0.0040	0.0194
General health	1.0000	-0.5417	0.0107	Education level (1)	0.0416	0.0030	0.0165
Politics scale	1.0000	0.0617	0.0042	Supervising other employees	0.0394	-0.0016	0.0088
Religious	1.0000	0.0884	0.0043	Unlimited job contract	0.0324	-0.0012	0.0076
Protestant	1.0000	0.4307	0.0262	Worked in a political party	0.0313	-0.0028	0.0180
Easternorthodox	1.0000	-0.4535	0.0404	Education level (4)	0.0261	-0.0027	0.0182
Jewish	1.0000	0.4382	0.0603	Living with children	0.0229	-0.0009	0.0064
Socializing	1.0000	0.1008	0.0061	Left political orientation	0.0183	-0.0009	0.0072
Intimate	1.0000	0.3710	0.0303	Daily activities limitation	0.0168	-0.0006	0.0057
Social activities	1.0000	0.1117	0.0098	Voted last election	0.0134	0.0005	0.0050
Trade union member	1.0000	0.1214	0.0213	Attending religious services	0.0131	0.0001	0.0014
Retired	0.9999	0.1926	0.0391	Doing other things	0.0124	-0.0012	0.0132
Pray	0.9998	0.0248	0.0045	Number of household members	0.0120	-0.0001	0.0014
Work hours	0.9996	-0.0033	0.0007	Right political orientation	0.0079	-0.0002	0.0035
Volunteer work	0.9981	0.1123	0.0235	Widowed	0.0069	0.0001	0.0037
Education level (5)	0.9871	-0.1163	0.0296	Education level (2)	0.0064	0.0001	0.0045
Other Christian	0.9150	0.2128	0.0873	Belong to a religion	0.0057	-0.0001	0.0054
Muslim	0.8853	-0.1677	0.0755	Education level (other)	0.0056	0.0001	0.0121
Family-owned business	0.8576	0.2320	0.1153	Education level (3)	0.0056	0.0001	0.0039
Education job	0.7064	0.1175	0.0853	Married	0.0054	0.0000	0.0025
Country born	0.7049	-0.0764	0.0557	Minority ethnic group	0.0051	-0.0001	0.0035
Age	0.6987	0.0022	0.0016	Community or military job	0.0046	0.0004	0.0123
Years of study (squared)	0.5349	-0.0007	0.0007	Other non-Christian	0.0031	0.0001	0.0072
Years of study	0.5136	0.0172	0.0179				

Table 3.5: Bayesian Model Averaging Analysis: Wave 4

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Discriminated	1.0000	-0.4877	0.0350	Interested in politics	0.4257	0.0145	0.0182
Income	1.0000	0.0408	0.0041	Gender	0.3531	-0.0199	0.0290
Borrow money	1.0000	0.1139	0.0073	Improved skills	0.3510	0.0230	0.0337
Income feeling	1.0000	-0.6571	0.0127	Self-employed	0.2872	0.0236	0.0400
Partner	1.0000	0.2557	0.0204	Eastern religion	0.2356	0.0825	0.1600
Separated	1.0000	-0.4052	0.0668	Divorced	0.0968	-0.0083	0.0275
Unemployed	1.0000	-0.6014	0.0468	Member of a political party	0.0769	0.0067	0.0257
Paid work	1.0000	-0.2265	0.0334	Work form house	0.0552	0.0027	0.0124
General health	1.0000	-0.5417	0.0107	Roman Catholic	0.0507	0.0042	0.0199
Politics scale	1.0000	0.0617	0.0042	Supervising other employees	0.0423	-0.0017	0.0090
Religious	1.0000	0.0884	0.0043	Education level (1)	0.0390	0.0028	0.0158
Pray	1.0000	0.0248	0.0045	Worked in a political party	0.0316	-0.0029	0.0182
Protestant	1.0000	0.4310	0.0265	Unlimited job contract	0.0314	-0.0012	0.0074
Easternorthodox	1.0000	-0.4533	0.0406	Education level (4)	0.0314	-0.0032	0.0199
Jewish	1.0000	0.4390	0.0604	Living with children	0.0284	-0.0011	0.0071
Socializing	1.0000	0.1007	0.0061	Left political orientation	0.0198	-0.0009	0.0075
Intimate	1.0000	0.3709	0.0303	Daily activities limitation	0.0171	-0.0006	0.0057
Social activities	1.0000	0.1118	0.0098	Voted last election	0.0138	0.0005	0.0051
Trade union member	1.0000	0.1216	0.0213	Attending religious services	0.0131	0.0001	0.0014
Retired	0.9999	0.1934	0.0394	Number of household members	0.0128	-0.0001	0.0014
Victim of burglary	0.9998	-0.1156	0.0208	Doing other things	0.0124	-0.0012	0.0131
Work hours	0.9996	-0.0033	0.0007	Widowed	0.0076	0.0002	0.0040
Volunteering work	0.9982	0.1123	0.0234	Right political orientation	0.0075	-0.0002	0.0034
Education level (5)	0.9892	-0.1174	0.0291	Belong to a religion	0.0057	-0.0001	0.0055
Other Christian	0.9162	0.2132	0.0872	Education level (3)	0.0056	0.0000	0.0035
Muslim	0.8839	-0.1680	0.0759	Married	0.0053	0.0000	0.0025
Family-owned business	0.8589	0.2326	0.1151	Other non-Christian	0.0049	0.0002	0.0092
Country born	0.7003	-0.0758	0.0558	Education level (2)	0.0048	0.0001	0.0038
Education job	0.6947	0.1150	0.0857	Minority ethnic group	0.0047	-0.0001	0.0033
Age	0.6797	0.0021	0.0016	Community or military job	0.0044	0.0004	0.0122
Years of study (squared)	0.5099	-0.0007	0.0007	Education level (other)	0.0035	0.0000	0.0095
Years of study	0.4868	0.0163	0.0178				

Table 3.6: Bayesian Model Averaging Analysis: Waves 1-3

Variable	PIP	Post Mean	Post SD	Variable	PIP	Post Mean	Post SD
(i)	(ii)	(iii)	(iv)	(v)	(vi)		
Discriminated	1.0000	-0.4957	0.0408	Worked in a political party	0.1464	-0.0210	0.0551
Income	1.0000	0.0659	0.0050	Voted last election	0.1125	0.0077	0.0234
Borrow money	1.0000	0.1063	0.0084	Country born	0.1032	-0.0097	0.0311
Income feeling	1.0000	-0.6315	0.0148	Right political orientation	0.0901	-0.0068	0.0237
Partner	1.0000	0.2443	0.0233	Daily activities limitation	0.0594	-0.0036	0.0157
Separated	1.0000	-0.4936	0.0775	Education level (3)	0.0593	-0.0034	0.0148
Unemployed	1.0000	-0.6020	0.0526	Age	0.0540	0.0001	0.0005
Retired	1.0000	0.2587	0.0373	Living with children	0.0529	-0.0026	0.0120
Paid work	1.0000	-0.2605	0.0345	Eastern religion	0.0495	0.0142	0.0692
General health	1.0000	-0.5317	0.0126	Divorced	0.0389	-0.0031	0.0172
Politics scale	1.0000	0.0691	0.0064	Work from house	0.0380	0.0019	0.0109
Religious	1.0000	0.0880	0.0050	Supervising other employees	0.0248	-0.0010	0.0072
Protestant	1.0000	0.3814	0.0236	Minority ethnic group	0.0188	-0.0016	0.0135
Easternorthodox	1.0000	-0.3089	0.0441	Left political orientation	0.0183	-0.0009	0.0080
Socializing	1.0000	0.0866	0.0070	Number of household members	0.0170	-0.0002	0.0019
Intimate	1.0000	0.3014	0.0358	Jewish	0.0149	0.0020	0.0194
Social activities	1.0000	0.1288	0.0112	Attending religious services	0.0140	0.0002	0.0017
Pray	0.9995	0.0269	0.0052	Education level (1)	0.0131	0.0006	0.0071
Victim of burglary	0.9968	-0.1133	0.0247	Interested in politics	0.0122	0.0002	0.0022
Volunteering work	0.9897	0.1165	0.0289	Unlimited job contract	0.0113	-0.0003	0.0036
Gender	0.9737	-0.0883	0.0262	Other non-Christian	0.0110	0.0020	0.0243
Trade union member	0.9708	0.0964	0.0287	Self-employed	0.0109	0.0004	0.0054
Education level (5)	0.9395	-0.1271	0.0481	Education level (2)	0.0104	0.0003	0.0044
Work hours	0.8534	-0.0022	0.0011	Roman Catholic	0.0094	0.0001	0.0033
Family-owned business	0.6979	0.2031	0.1506	Community or military job	0.0072	0.0011	0.0221
Years of study (squared)	0.5090	-0.0004	0.0005	Belong to a religion	0.0068	0.0002	0.0047
Other Christian	0.4807	0.0988	0.1118	Doing other things	0.0066	-0.0002	0.0067
Education job	0.2770	0.0413	0.0720	Education level (4)	0.0064	0.0002	0.0057
Improved skills	0.2472	0.0174	0.0326	Muslim	0.0058	0.0000	0.0045
Years of study	0.1957	0.0051	0.0116	Married	0.0057	0.0001	0.0031
Member of a political party	0.1885	0.0235	0.0531	Widowed	0.0054	0.0001	0.0034

Table 3.7: Bayesian Model Averaging Analysis: Waves 1-4

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Discriminated	1.0000	-0.4877	0.0350	Interested in politics	0.4313	0.0148	0.0184
Income	1.0000	0.0408	0.0041	Improved skills	0.3632	0.0239	0.0341
Borrow money	1.0000	0.1139	0.0073	Gender	0.3477	-0.0196	0.0289
Income feeling	1.0000	-0.6570	0.0127	Self-employed	0.2858	0.0235	0.0400
Partner	1.0000	0.2557	0.0203	Eastern religion	0.2486	0.0874	0.1638
Separated	1.0000	-0.4054	0.0668	Divorced	0.0982	-0.0085	0.0278
Unemployed	1.0000	-0.6010	0.0468	Member of a political party	0.0805	0.0069	0.0257
Paid work	1.0000	-0.2261	0.0333	Roman catholic	0.0601	0.0051	0.0219
General health	1.0000	-0.5418	0.0106	Work from house	0.0580	0.0029	0.0127
Politics scale	1.0000	0.0617	0.0042	Education level (1)	0.0445	0.0033	0.0180
Religious	1.0000	0.0884	0.0043	Supervising other employees	0.0413	-0.0017	0.0090
Protestant	1.0000	0.4315	0.0277	Education level (4)	0.0316	-0.0032	0.0199
Easternorthodox	1.0000	-0.4523	0.0415	Unlimited contract	0.0285	-0.0011	0.0071
Jewish	1.0000	0.4391	0.0611	Worked in a political part	0.0276	-0.0025	0.0170
Socializing	1.0000	0.1008	0.0061	Living with children	0.0265	-0.0010	0.0067
Intimate	1.0000	0.3710	0.0303	Left political orientation	0.0190	-0.0009	0.0073
Social activities	1.0000	0.1116	0.0098	Daily activities limitation	0.0165	-0.0006	0.0056
Trade union member	1.0000	0.1213	0.0213	Voted last election	0.0145	0.0005	0.0052
Victim of burglary	0.9999	-0.1155	0.0208	Attending religious services	0.0144	0.0002	0.0015
Retired	0.9998	0.1922	0.0392	Doing other things	0.0128	-0.0012	0.0134
Pray	0.9998	0.0248	0.0045	Number of household members	0.0124	-0.0001	0.0014
Work hours	0.9996	-0.0033	0.0007	Education level (3)	0.0076	0.0002	0.0054
Volunteering work	0.9983	0.1121	0.0234	Right political orientation	0.0076	-0.0002	0.0035
Education level (5)	0.9843	-0.1154	0.0303	Education level (2)	0.0068	0.0003	0.0063
Other Christian	0.9156	0.2141	0.0880	Belong to a religion	0.0064	-0.0002	0.0065
Muslim	0.8689	-0.1642	0.0781	Education level (other)	0.0061	0.0001	0.0127
Family-owned business	0.8561	0.2316	0.1156	Minority ethnic group	0.0061	-0.0001	0.0038
Education job	0.7078	0.1181	0.0854	Married	0.0060	0.0000	0.0026
Age	0.7045	0.0022	0.0016	Widowed	0.0053	0.0001	0.0033
Country born	0.6991	-0.0757	0.0558	Other non-Christian	0.0047	0.0002	0.0088
Years of study (squared)	0.5673	-0.0008	0.0007	Community or military job	0.0039	0.0003	0.0115
Years of study	0.5449	0.0183	0.0179				

Table 3.8: Bayesian Model Averaging Analysis: Wave 1 (No Subjective Variables)

Variable	PIP	Post Mean	Post SD	Variable	PIP	Post Mean	Post SD
	(i)	(ii)	(iii)		(iv)	(v)	(vi)
Age	1.0000	0.0068	0.0009	Family-owned business	0.9767	0.3051	0.0863
Discriminated	1.0000	-0.6010	0.0356	Improved skills	0.9608	0.0855	0.0273
Income	1.0000	0.1131	0.0040	Gender	0.9214	-0.0679	0.0266
Borrow money	1.0000	0.1980	0.0073	Education job	0.9206	0.1644	0.0649
Victim of burglary	1.0000	-0.1634	0.0212	Education level (5)	0.9114	-0.0932	0.0377
Partner	1.0000	0.2387	0.0255	Country born	0.7593	-0.0845	0.0546
Divorced	1.0000	-0.2107	0.0364	Self-employed	0.5897	0.0555	0.0509
Separated	1.0000	-0.5487	0.0693	Education level (1)	0.4737	0.0529	0.0642
Unemployed	1.0000	-0.7858	0.0464	Widowed	0.1278	-0.0134	0.0378
Retired	1.0000	0.2100	0.0356	Work from house	0.1014	0.0061	0.0195
General health	1.0000	-0.6161	0.0109	Pray	0.0930	-0.0011	0.0036
Politics scale	1.0000	0.0722	0.0043	Member of a political party	0.0638	0.0053	0.0226
Roman Catholic	1.0000	0.2345	0.0291	Education level (4)	0.0622	-0.0078	0.0330
Protestant	1.0000	0.6637	0.0317	Daily activities limitation	0.0506	-0.0027	0.0127
Easternorthodox	1.0000	-0.4732	0.0417	Education level (3)	0.0459	0.0042	0.0219
Other Christian	1.0000	0.4716	0.0634	Education level (2)	0.0452	0.0051	0.0260
Jewish	1.0000	0.4428	0.0603	Other non-Christian	0.0407	0.0106	0.0569
Years of study	1.0000	0.0579	0.0107	Left political orientation	0.0261	-0.0013	0.0092
Years of study (squared)	1.0000	-0.0020	0.0004	Worked in a political party	0.0252	-0.0022	0.0160
Socializing	1.0000	0.1157	0.0061	Unlimited job contract	0.0252	-0.0010	0.0072
Intimate	1.0000	0.3993	0.0305	Voted last election	0.0180	0.0008	0.0064
Social activities	1.0000	0.1265	0.0099	Belong to a religion	0.0124	0.0011	0.0119
Attending religious services	0.9999	-0.0413	0.0066	Doing other things	0.0106	-0.0010	0.0120
Work hours	0.9999	-0.0035	0.0007	Community or military job	0.0073	0.0013	0.0206
Trade union member	0.9999	0.1185	0.0218	Education level (other)	0.0060	0.0006	0.0154
Paid work	0.9998	-0.2026	0.0317	Married	0.0044	0.0000	0.0024
Volunteer work	0.9992	0.1175	0.0236	Right political orientation	0.0043	-0.0001	0.0022
Number of household members	0.9986	-0.0350	0.0073	Muslim	0.0039	0.0001	0.0039
Eastern religion	0.9984	0.5847	0.1237				

Table 3.9: Bayesian Model Averaging Analysis: Wave 2 (No Subjective Variables)

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Age	1.0000	0.0068	0.0009	Family-owned business	0.9740	0.3040	0.0876
Discriminated	1.0000	-0.6010	0.0356	Improved skills	0.9586	0.0853	0.0275
Income	1.0000	0.1131	0.0040	Education level (5)	0.9336	-0.0955	0.0352
Borrow money	1.0000	0.1980	0.0073	Education job	0.9185	0.1640	0.0653
Victim of burglary	1.0000	-0.1633	0.0212	Gender	0.9129	-0.0672	0.0272
Partner	1.0000	0.2388	0.0254	Country born	0.7559	-0.0841	0.0547
Divorced	1.0000	-0.2106	0.0364	Self-employed	0.5719	0.0538	0.0510
Separated	1.0000	-0.5486	0.0692	Education level (1)	0.4681	0.0501	0.0600
Unemployed	1.0000	-0.7861	0.0465	Widowed	0.1234	-0.0129	0.0372
Retired	1.0000	0.2098	0.0356	Work from house	0.1042	0.0062	0.0197
Paid work	1.0000	-0.2026	0.0317	Pray	0.0978	-0.0011	0.0037
Work hours	1.0000	-0.0035	0.0007	Member of a political party	0.0644	0.0054	0.0229
General health	1.0000	-0.6161	0.011	Education level (4)	0.0623	-0.0078	0.0331
Politics scale	1.0000	0.0722	0.0043	Daily activities limitation	0.0565	-0.0030	0.0134
Roman Catholic	1.0000	0.2344	0.0291	Other non-Christian	0.0405	0.0105	0.0568
Protestant	1.0000	0.6636	0.0317	Unlimited job contract	0.0301	-0.0012	0.0079
Easternorthodox	1.0000	-0.4731	0.0417	Education level (3)	0.0276	0.0021	0.0160
Other Christian	1.0000	0.4715	0.0634	Worked in a political party	0.0257	-0.0023	0.0164
Jewish	1.0000	0.4430	0.0603	Education level (2)	0.0252	0.0026	0.0188
Years of study	1.0000	0.0578	0.0106	Left political orientation	0.0246	-0.0012	0.0089
Years of study (squared)	1.0000	-0.0020	0.0004	Voted last election	0.0188	0.0008	0.0066
Socializing	1.0000	0.1157	0.0061	Belong to a religion	0.0141	0.0013	0.0126
Intimate	1.0000	0.3994	0.0305	Doing other things	0.0102	-0.0009	0.0117
Social activities	1.0000	0.1265	0.0099	Community or military job	0.0082	0.0014	0.0219
Trade union member	1.0000	0.1183	0.0217	Muslim	0.0051	0.0001	0.0044
Attending religious services	0.9995	-0.0413	0.0067	Right political orientation	0.0050	-0.0001	0.0024
Volunteer work	0.9992	0.1175	0.0236	Married	0.0050	0.0000	0.0026
Eastern religion	0.9983	0.5848	0.1238	Education level (other)	0.0045	0.0004	0.0126
Number of household members	0.9982	-0.0350	0.0074				

Table 3.10: Bayesian Model Averaging Analysis: Wave 3 (No Subjective Variables)

Variable	PIP	Post Mean	Post SD	Variable	PIP	Post Mean	Post SD
	(i)	(ii)	(iii)		(iv)	(v)	(vi)
Age	1.0000	0.0068	0.0009	Family-owned business	0.9754	0.3044	0.0869
Discriminated	1.0000	-0.6011	0.0356	Improved skills	0.9556	0.0851	0.0279
Income	1.0000	0.1131	0.004	Gender	0.9213	-0.0679	0.0266
Borrow money	1.0000	0.1981	0.0073	Education level (5)	0.9211	-0.0943	0.0368
Victim of burglary	1.0000	-0.1633	0.0212	Education job	0.9144	0.1634	0.0661
Partner	1.0000	0.2385	0.0256	Country born	0.7646	-0.0851	0.0543
Divorced	1.0000	-0.2109	0.0365	Self-employed	0.5702	0.0536	0.0510
Separated	1.0000	-0.5489	0.0693	Education level (1)	0.4702	0.0513	0.062
Unemployed	1.0000	-0.7863	0.0465	Widowed	0.1318	-0.0139	0.0384
Retired	1.0000	0.2098	0.0356	Pray	0.0992	-0.0011	0.0037
Paid work	1.0000	-0.2026	0.0318	Work from house	0.0979	0.0058	0.0191
Work hours	1.0000	-0.0035	0.0007	Member of a political party	0.0647	0.0055	0.0229
General health	1.0000	-0.6161	0.0110	Education level (4)	0.0640	-0.0080	0.0337
Politics scale	1.0000	0.0722	0.0044	Daily activities limitation	0.0529	-0.0028	0.0130
Roman Catholic	1.0000	0.2345	0.0292	Other non-Christian	0.0438	0.0114	0.0590
Protestant	1.0000	0.6638	0.0318	Education level (3)	0.0371	0.0030	0.0191
Easternorthodox	1.0000	-0.4732	0.0417	Education level (2)	0.0346	0.0037	0.0225
Other Christian	1.0000	0.4716	0.0634	Unlimited job contract	0.0312	-0.0013	0.0081
Jewish	1.0000	0.4427	0.0603	Worked in a political party	0.0263	-0.0024	0.0165
Years of study	1.0000	0.0578	0.0107	Left political orientation	0.0251	-0.0013	0.0090
Years of study (squared)	1.0000	-0.0020	0.0004	Voted last election	0.0216	0.0009	0.0070
Socializing	1.0000	0.1157	0.0061	Belong to a religion	0.0147	0.0013	0.0128
Intimate	1.0000	0.3994	0.0305	Doing other things	0.0117	-0.0011	0.0126
Social activities	1.0000	0.1265	0.0099	Community or military job	0.0088	0.0015	0.0226
Trade union member	1.0000	0.1183	0.0217	Right political orientation	0.0066	-0.0001	0.0027
Attending religious services	0.9998	-0.0413	0.0066	Education level (other)	0.0055	0.0005	0.0143
Volunteer work	0.9991	0.1176	0.0236	Muslim	0.0041	0.0001	0.0039
Number of household members	0.9979	-0.0350	0.0074	Married	0.0040	0.0000	0.0023
Eastern religion	0.9974	0.5841	0.1250				

Table 3.11: Bayesian Model Averaging Analysis: Wave 4 (No Subjective Variables)

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Age	1.0000	0.0068	0.0009	Family-owned business	0.9744	0.3041	0.0874
Discriminated	1.0000	-0.6010	0.0356	Improved skills	0.9579	0.0852	0.0276
Income	1.0000	0.1131	0.004	Education level (5)	0.9301	-0.0952	0.0356
Borrow money	1.0000	0.1980	0.0073	Education job	0.9194	0.1642	0.0652
Victim of burglary	1.0000	-0.1633	0.0212	Gender	0.9177	-0.0676	0.0269
Partner	1.0000	0.2384	0.0256	Country born	0.7569	-0.0842	0.0547
Divorced	1.0000	-0.2109	0.0365	Self-employed	0.5725	0.0538	0.0510
Separated	1.0000	-0.5489	0.0693	Education level (1)	0.4667	0.0502	0.0605
Unemployed	1.0000	-0.7860	0.0464	Widowed	0.1339	-0.0141	0.0386
Retired	1.0000	0.2099	0.0356	Work from house	0.1040	0.0062	0.0197
Paid work	1.0000	-0.2024	0.0317	Pray	0.0972	-0.0011	0.0037
Work hours	1.0000	-0.0035	0.0007	Member of a political party	0.0700	0.0059	0.0237
General health	1.0000	-0.6161	0.0110	Education level (4)	0.0644	-0.0081	0.0336
Politics scale	1.0000	0.0722	0.0043	Daily activities limitation	0.0520	-0.0027	0.0129
Roman Catholic	1.0000	0.2344	0.0291	Other non-Christian	0.0432	0.0112	0.0585
Protestant	1.0000	0.6636	0.0317	Education level (3)	0.0301	0.0023	0.0168
Easternorthodox	1.0000	-0.4732	0.0417	Unlimited job contract	0.0293	-0.0012	0.0078
Other Christian	1.0000	0.4715	0.0634	Education level (2)	0.0276	0.0029	0.0198
Jewish	1.0000	0.4429	0.0604	Worked in a political party	0.0270	-0.0024	0.0168
Years of study	1.0000	0.0578	0.0106	Left political orientation	0.0228	-0.0012	0.0086
Years of study (squared)	1.0000	-0.0020	0.0004	Voted last election	0.0195	0.0008	0.0067
Socializing	1.0000	0.1157	0.0061	Belong to a religion	0.0140	0.0012	0.0125
Intimate	1.0000	0.3993	0.0305	Doing other things	0.0116	-0.0011	0.0125
Social activities	1.0000	0.1264	0.0099	Community or military job	0.0086	0.0015	0.0224
Trade union member	1.0000	0.1183	0.0217	Muslim	0.0063	0.0001	0.0048
Attending religious services	0.9998	-0.0413	0.0066	Right political orientation	0.0053	-0.0001	0.0025
Volunteer work	0.9995	0.1175	0.0235	Education level (other)	0.0039	0.0004	0.0120
Eastern religion	0.9983	0.5848	0.1238	Married	0.0034	0.0000	0.0021
Number of household members	0.9976	-0.0350	0.0074				

Table 3.12: Bayesian Model Averaging Analysis: Wave 1-3 (No Subjective Variables)

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Discriminated	1.0000	-0.6024	0.0421	Family-owned business	0.8580	0.2736	0.1364
Income	1.0000	0.1411	0.0049	Eastern religion	0.8080	0.4225	0.2420
Borrow money	1.0000	0.1855	0.0084	Trade union member	0.7934	0.0704	0.0420
Victim of burglary	1.0000	-0.1603	0.0244	Voted last election	0.7425	0.0705	0.0476
Partner	1.0000	0.2291	0.0286	Improved skills	0.5131	0.0424	0.0451
Separated	1.0000	-0.6387	0.0796	Education level (5)	0.4098	-0.0473	0.0640
Unemployed	1.0000	-0.7616	0.0541	Education level (3)	0.1814	-0.0149	0.0341
Retired	1.0000	0.2586	0.0411	Daily activities limitation	0.1477	-0.0107	0.0278
Paid work	1.0000	-0.2209	0.0368	Other non-Christian	0.1466	0.0568	0.1481
General health	1.0000	-0.5948	0.0140	Education level (1)	0.1262	0.0125	0.0361
Politics scale	1.0000	0.0791	0.0064	Member of a political party	0.1041	0.0120	0.0387
Protestant	1.0000	0.5757	0.0548	Worked in a political party	0.0972	-0.0130	0.0436
Years of study	1.0000	0.0544	0.0110	Belong to a religion	0.0879	0.0169	0.0601
Years of study (squared)	1.0000	-0.0022	0.0004	Right political orientation	0.0625	-0.0046	0.0197
Socializing	1.0000	0.0967	0.0072	Country born	0.0563	-0.0049	0.0223
Intimate	1.0000	0.3127	0.0365	Widowed	0.0511	-0.0051	0.0243
Social activities	1.0000	0.1448	0.0115	Work from house	0.0505	0.0028	0.0136
Gender	0.9992	-0.1093	0.0225	Left political orientation	0.0433	-0.0028	0.0149
Age	0.9990	0.0053	0.0011	Pray	0.0361	-0.0004	0.0024
Divorced	0.9961	-0.1897	0.0432	Education level (2)	0.0290	0.0017	0.0127
Attending religious services	0.9949	-0.0383	0.0078	Unlimited job contract	0.0175	-0.0006	0.0054
Other Christian	0.9943	0.4372	0.0875	Jewish	0.0116	0.0015	0.0178
Easternorthodox	0.9921	-0.2470	0.0696	Self-employed	0.0108	0.0005	0.0057
Volunteer work	0.9809	0.1147	0.0314	Community or military job	0.0103	0.0026	0.0339
Roman Catholic	0.9759	0.2242	0.0557	Education level (4)	0.0084	0.0003	0.0079
Number of household members	0.9455	-0.0320	0.0112	Doing other things	0.0062	-0.0001	0.0065
Muslim	0.9186	0.2523	0.0994	Education level (other)	0.0059	0.0006	0.0153
Work hours	0.8941	-0.0024	0.0010	Married	0.0044	0.0000	0.0027
Education job	0.8920	0.1828	0.0812				

Table 3.13: Bayesian Model Averaging Analysis: Wave 1-4 (No Subjective Variables)

Variable	PIP	Post Mean	Post SD	Variable	PIP	Post Mean	Post SD
	(i)	(ii)	(iii)		(iv)	(v)	(vi)
Age	1.0000	0.0068	0.0009	Family-owned business	0.9765	0.3049	0.0863
Discriminated	1.0000	-0.6011	0.0356	Improved skills	0.9596	0.0854	0.0274
Income	1.0000	0.1131	0.004	Education level (5)	0.9233	-0.0945	0.0364
Borrow money	1.0000	0.198	0.0073	Gender	0.9214	-0.0679	0.0266
Victim of burglary	1.0000	-0.1634	0.0212	Education job	0.9144	0.1634	0.0661
Partner	1.0000	0.2383	0.0258	Country born	0.7656	-0.0853	0.0543
Divorced	1.0000	-0.2111	0.0366	Self-employed	0.5800	0.0546	0.0510
Separated	1.0000	-0.5491	0.0693	Education level (1)	0.4616	0.0501	0.0612
Unemployed	1.0000	-0.7862	0.0465	Widowed	0.1362	-0.0143	0.0390
Retired	1.0000	0.2099	0.0356	Work from house	0.1040	0.0062	0.0197
Paid work	1.0000	-0.2027	0.0317	Pray	0.0972	-0.0011	0.0036
General health	1.0000	-0.6161	0.0110	Member of a political party	0.0691	0.0058	0.0236
Politics scale	1.0000	0.0722	0.0044	Education level (4)	0.0672	-0.0084	0.0343
Roman Catholic	1.0000	0.2345	0.0291	Daily activities limitation	0.0561	-0.0030	0.0134
Protestant	1.0000	0.6637	0.0317	Other non-Christian	0.0434	0.0113	0.0587
Easternorthodox	1.0000	-0.4732	0.0417	Education level (3)	0.0341	0.0027	0.0179
Other Christian	1.0000	0.4716	0.0634	Education level (2)	0.0311	0.0033	0.0212
Jewish	1.0000	0.4426	0.0603	Unlimited job contract	0.0287	-0.0012	0.0078
Years of study	1.0000	0.0577	0.0106	Left political orientation	0.0275	-0.0014	0.0095
Years of study (squared)	1.0000	-0.0020	0.0004	Worked in a political party	0.0266	-0.0024	0.0169
Socializing	1.0000	0.1157	0.0061	Voted last election	0.0221	0.0009	0.0071
Intimate	1.0000	0.3993	0.0305	Belong to a religion	0.0129	0.0012	0.0121
Social activities	1.0000	0.1264	0.0099	Doing other things	0.0096	-0.0009	0.0114
Work hours	0.9999	-0.0035	0.0007	Community or military job	0.0071	0.0012	0.0202
Trade union member	0.9999	0.1184	0.0217	Education level (other)	0.0066	0.0006	0.0152
Volunteer work	0.9998	0.1176	0.0234	Muslim	0.0061	0.0001	0.0048
Attending religious services	0.9997	-0.0413	0.0066	Right political orientation	0.0056	-0.0001	0.0025
Number of household members	0.9983	-0.0350	0.0074	Married	0.0045	0.0000	0.0025
Eastern religion	0.9983	0.5844	0.1238				

Table 3.14: Jointness of Happiness Determinants: Wave 1

Variable	$\mathbf{x_h/x_l}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Gender	1	.	-0.8	-0.9	-6	-0.7	-0.7	-0.7	-0.7	-0.7	-5.2	-0.7	-6.1	-2.3	-0.7	-6	-3.6
Age	2	-0.8	.	0.1	-6	0.8	0.8	0.8	0.8	0.8	-5.1	0.8	-6	-2.2	0.8	-5.7	-4.3
Country born	3	-0.9	0.1	.	-6.1	0.9	0.9	0.9	0.9	0.9	-4.8	0.9	-5.7	-2.3	0.9	-5.8	-3.8
Minority ethnic group	4	-6	-6	-6.1	.	-6.2	-6.2	-6.2	-6.2	-6.2	PS	-6.2	PS	-6.1	-6.2	PS	PS
Discriminated	5	-0.7	0.8	0.9	-6.2	.	PC	PC	PC	PC	-4.8	PC	-5.8	-2.3	PC	-5.9	-3.8
Income	6	-0.7	0.8	0.9	-6.2	PC	.	PC	PC	PC	-4.8	PC	-5.8	-2.3	PC	-5.9	-3.8
Borrow money	7	-0.7	0.8	0.9	-6.2	PC	PC	.	PC	PC	-4.8	PC	-5.8	-2.3	PC	-5.9	-3.8
Victim of burglary	8	-0.7	0.8	0.9	-6.2	PC	PC	PC	.	PC	-4.8	PC	-5.8	-2.3	PC	-5.9	-3.8
Income feeling	9	-0.7	0.8	0.9	-6.2	PC	PC	PC	PC	.	-4.8	PC	-5.8	-2.3	PC	-5.9	-3.8
Number of household members	10	-5.2	-5.1	-4.8	PS	-4.8	-4.8	-4.8	-4.8	-4.8	.	-4.8	PS	-5.8	-4.8	PS	PS
Partner	11	-0.7	0.8	0.9	-6.2	PC	PC	PC	PC	PC	-4.8	.	-5.8	-2.3	PC	-5.9	-3.8
Widowed	12	-6.1	-6	-5.7	PS	-5.8	-5.8	-5.8	-5.8	-5.8	PS	-5.8	.	-6.7	-5.8	PS	PS
Divorced	13	-2.3	-2.2	-2.3	-6.1	-2.3	-2.3	-2.3	-2.3	-2.3	-5.8	-2.3	-6.7	.	-2.3	-8.6	-4.6
Separated	14	-0.7	0.8	0.9	-6.2	PC	PC	PC	PC	PC	-4.8	PC	-5.8	-2.3	.	-5.9	-3.8
Married	15	-6	-5.7	-5.8	PS	-5.9	-5.9	-5.9	-5.9	-5.9	PS	-5.9	PS	-8.6	-5.9	.	PS
Living with children	16	-3.6	-4.3	-3.8	PS	-3.8	-3.8	-3.8	-3.8	-3.8	PS	-3.8	PS	-4.6	-3.8	PS	.
Unemployed	17	-0.7	0.8	0.9	-6.2	PC	PC	PC	PC	PC	-4.8	PC	-5.8	-2.3	PC	-5.9	-3.8
Retired	18	-0.7	0.8	0.9	-6.2	10.6	10.6	10.6	10.6	10.6	-4.8	10.6	-5.8	-2.3	10.6	-5.9	-3.8
Paid work	19	-0.7	0.8	0.9	-6.2	PC	PC	PC	PC	PC	-4.8	PC	-5.8	-2.3	PC	-5.9	-3.8
Community or military job	20	-5.5	-5.4	-5.5	PS	-5.6	-5.6	-5.6	-5.6	-5.6	PS	-5.6	PS	-5.9	-5.6	PS	PS
Work from house	21	-4.3	-2.9	-3	-7.3	-3	-3	-3	-3	-3	-5.5	-3	PS	-3.7	-3	-7	-4.4
Doing other things	22	-4.7	-4.7	-4.6	PS	-4.7	-4.7	-4.7	-4.7	-4.7	PS	-4.7	PS	-5	-4.7	PS	-8.1
Education job	23	-0.7	1.4	0.1	-6	0.8	0.8	0.8	0.8	0.8	-4.8	0.8	-5.8	-2.3	0.8	-5.7	-4.2
Self-employed	24	-1.4	-1.5	-1	-6.4	-0.9	-0.9	-0.9	-0.9	-0.9	-4.8	-0.9	-6.3	-2.6	-0.9	-6.3	-3.8
Family-owned business	25	-0.8	0.4	0.7	-6.2	1.9	1.9	1.9	1.9	1.9	-4.7	1.9	-5.8	-2.4	1.9	-5.8	-3.8
Unlimited contract	26	-3.7	-3.7	-3.6	-7.2	-3.7	-3.7	-3.7	-3.7	-3.7	-9.2	-3.7	-6.6	-4.4	-3.7	-6.4	-5.4
Work hours	27	-0.7	0.8	0.9	-6.2	10.5	10.5	10.5	10.5	10.5	-4.8	10.5	-5.8	-2.3	10.5	-5.9	-3.8
Supervising other employees	28	-4	-3.3	-3.3	-7.1	-3.3	-3.3	-3.3	-3.3	-3.3	-5.7	-3.3	PS	-4.2	-3.3	PS	-5.5
General health	29	-0.7	0.8	0.9	-6.2	PC	PC	PC	PC	PC	-4.8	PC	-5.8	-2.3	PC	-5.9	-3.8
Daily activities limitation	30	-4.8	-4.5	-4.5	PS	-4.6	-4.6	-4.6	-4.6	-4.6	PS	-4.6	PS	-5	-4.6	PS	-6.2
Left political orientation	31	-4.3	-4.3	-4.3	PS	-4.3	-4.3	-4.3	-4.3	-4.3	PS	-4.3	-6.5	-5	-4.3	PS	-5
Right political orientation	32	-5.6	-5.5	-5.5	PS	-5.5	-5.5	-5.5	-5.5	-5.5	PS	-5.5	PS	-6.7	-5.5	PS	PS

Table 3.15: Jointness of Happiness Determinants: Wave 1 (Continued)

Variable	$\mathbf{x_h/x_l}$	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Gender	1	-0.7	-0.7	-0.7	-5.5	-4.3	-4.7	-0.7	-1.4	-0.8	-3.7	-0.7	-4	-0.7	-4.8	-4.3	-5.6
Age	2	0.8	0.8	0.8	-5.4	-2.9	-4.7	1.4	-1.5	0.4	-3.7	0.8	-3.3	0.8	-4.5	-4.3	-5.5
Country born	3	0.9	0.9	0.9	-5.5	-3	-4.6	0.1	-1	0.7	-3.6	0.9	-3.3	0.9	-4.5	-4.3	-5.5
Minority ethnic group	4	-6.2	-6.2	-6.2	PS	-7.3	PS	-6	-6.4	-6.2	-7.2	-6.2	-7.1	-6.2	PS	PS	PS
Discriminated	5	PC	10.6	PC	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	10.5	-3.3	PC	-4.6	-4.3	-5.5
Income	6	PC	10.6	PC	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	10.5	-3.3	PC	-4.6	-4.3	-5.5
Borrow money	7	PC	10.6	PC	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	10.5	-3.3	PC	-4.6	-4.3	-5.5
Victim of burglary	8	PC	10.6	PC	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	10.5	-3.3	PC	-4.6	-4.3	-5.5
Income feeling	9	PC	10.6	PC	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	10.5	-3.3	PC	-4.6	-4.3	-5.5
Number of household members	10	-4.8	-4.8	-4.8	PS	-5.5	PS	-4.8	-4.8	-4.7	-9.2	-4.8	-5.7	-4.8	PS	PS	PS
Partner	11	PC	10.6	PC	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	10.5	-3.3	PC	-4.6	-4.3	-5.5
Widowed	12	-5.8	-5.8	-5.8	PS	PS	PS	-5.8	-6.3	-5.8	-6.6	-5.8	PS	-5.8	PS	-6.5	PS
Divorced	13	-2.3	-2.3	-2.3	-5.9	-3.7	-5	-2.3	-2.6	-2.4	-4.4	-2.3	-4.2	-2.3	-5	-5	-6.7
Separated	14	PC	10.6	PC	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	10.5	-3.3	PC	-4.6	-4.3	-5.5
Married	15	-5.9	-5.9	-5.9	PS	-7	PS	-5.7	-6.3	-5.8	-6.4	-5.9	PS	-5.9	PS	PS	PS
Living with children	16	-3.8	-3.8	-3.8	PS	-4.4	-8.1	-4.2	-3.8	-3.8	-5.4	-3.8	-5.5	-3.8	-6.2	-5	PS
Unemployed	17	.	10.6	PC	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	10.5	-3.3	PC	-4.6	-4.3	-5.5
Retired	18	10.6	.	10.6	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	9.8	-3.3	10.6	-4.6	-4.3	-5.5
Paid work	19	PC	10.6	.	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	10.5	-3.3	PC	-4.6	-4.3	-5.5
Community or military job	20	-5.6	-5.6	-5.6	.	-7.5	PS	-5.4	-5.7	-5.5	PS	-5.6	-6.3	-5.6	PS	PS	PS
Work from house	21	-3	-3	-3	-7.5	.	-6	-2.9	-3.2	-3	-4.4	-3	-4.2	-3	-5.2	-5.5	-7.8
Doing other things	22	-4.7	-4.7	-4.7	PS	-6	.	-4.7	-4.7	-4.7	-5.9	-4.7	-5.5	-4.7	PS	-9.2	PS
Education job	23	0.8	0.8	0.8	-5.4	-2.9	-4.7	.	-1.4	0.5	-3.8	0.8	-3.4	0.8	-4.6	-4.3	-5.5
Self-employed	24	-0.9	-0.9	-0.9	-5.7	-3.2	-4.7	-1.4	.	-0.9	-6.3	-0.9	-3.3	-0.9	-4.6	-4.4	-5.5
Family-owned business	25	1.9	1.9	1.9	-5.5	-3	-4.7	0.5	-0.9	.	-3.7	1.9	-3.3	1.9	-4.5	-4.3	-5.5
Unlimited contract	26	-3.7	-3.7	-3.7	PS	-4.4	-5.9	-3.8	-6.3	-3.7	.	-3.7	-4.6	-3.7	-6.2	-4.5	-8.4
Work hours	27	10.5	9.8	10.5	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	.	-3.3	10.5	-4.6	-4.3	-5.5
Supervising other employees	28	-3.3	-3.3	-3.3	-6.3	-4.2	-5.5	-3.4	-3.3	-3.3	-4.6	-3.3	.	-3.3	-5.8	-6.3	-7.3
General health	29	PC	10.6	PC	-5.6	-3	-4.7	0.8	-0.9	1.9	-3.7	10.5	-3.3	.	-4.6	-4.3	-5.5
Daily activities limitation	30	-4.6	-4.6	-4.6	PS	-5.2	PS	-4.6	-4.6	-4.5	-6.2	-4.6	-5.8	-4.6	.	-8.6	PS
Left political orientation	31	-4.3	-4.3	-4.3	PS	-5.5	-9.2	-4.3	-4.4	-4.3	-4.5	-4.3	-6.3	-4.3	-8.6	.	PS
Right political orientation	32	-5.5	-5.5	-5.5	PS	-7.8	PS	-5.5	-5.5	-5.5	-8.4	-5.5	-7.3	-5.5	PS	PS	.

Table 3.16: Jointness of Happiness Determinants: Wave 2

Variable	x_h/x_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Gender	1	.	-0.8	-0.9	-5.7	-0.7	-0.7	-0.7	-0.7	-0.7	-5	-0.7	-6.5	-2.3	-0.7	-6.2	-3.8
Age	2	-0.8	.	0.1	-5.8	0.9	0.9	0.9	0.9	0.9	-5.4	0.9	-6	-2.2	0.9	-6.3	-4.3
Country born	3	-0.9	0.1	.	-5.8	0.9	0.9	0.9	0.9	0.9	-4.9	0.9	-5.6	-2.4	0.9	-6.2	-3.9
Minority ethnic group	4	-5.7	-5.8	-5.8	.	-5.8	-5.8	-5.8	-5.8	-5.8	PS	-5.8	PS	-6.6	-5.8	PS	PS
Discriminated	5	-0.7	0.9	0.9	-5.8	.	PC	PC	PC	PC	-4.9	PC	-5.8	-2.4	PC	-6.3	-3.9
Income	6	-0.7	0.9	0.9	-5.8	PC	.	PC	PC	PC	-4.9	PC	-5.8	-2.4	PC	-6.3	-3.9
Borrow money	7	-0.7	0.9	0.9	-5.8	PC	PC	.	PC	PC	-4.9	PC	-5.8	-2.4	PC	-6.3	-3.9
Victim of burglary	8	-0.7	0.9	0.9	-5.8	PC	PC	PC	.	PC	-4.9	PC	-5.8	-2.4	PC	-6.3	-3.9
Income feeling	9	-0.7	0.9	0.9	-5.8	PC	PC	PC	PC	.	-4.9	PC	-5.8	-2.4	PC	-6.3	-3.9
Number of household members	10	-5	-5.4	-4.9	PS	-4.9	-4.9	-4.9	-4.9	-4.9	.	-4.9	PS	-5.4	-4.9	PS	PS
Partner	11	-0.7	0.9	0.9	-5.8	PC	PC	PC	PC	PC	-4.9	.	-5.8	-2.4	PC	-6.3	-3.9
Widowed	12	-6.5	-6	-5.6	PS	-5.8	-5.8	-5.8	-5.8	-5.8	PS	-5.8	.	-6	-5.8	PS	-6
Divorced	13	-2.3	-2.2	-2.4	-6.6	-2.4	-2.4	-2.4	-2.4	-2.4	-5.4	-2.4	-6	.	-2.4	-7.9	-5
Separated	14	-0.7	0.9	0.9	-5.8	PC	PC	PC	PC	PC	-4.9	PC	-5.8	-2.4	.	-6.3	-3.9
Married	15	-6.2	-6.3	-6.2	PS	-6.3	-6.3	-6.3	-6.3	-6.3	PS	-6.3	PS	-7.9	-6.3	.	PS
Living with children	16	-3.8	-4.3	-3.9	PS	-3.9	-3.9	-3.9	-3.9	-3.9	PS	-3.9	-6	-5	-3.9	PS	.
Unemployed	17	-0.7	0.9	0.9	-5.8	PC	PC	PC	PC	PC	-4.9	PC	-5.8	-2.4	PC	-6.3	-3.9
Retired	18	-0.7	0.9	0.9	-5.8	PC	PC	PC	PC	PC	-4.9	PC	-5.8	-2.4	PC	-6.3	-3.9
Paid work	19	-0.7	0.9	0.9	-5.8	PC	PC	PC	PC	PC	-4.9	PC	-5.8	-2.4	PC	-6.3	-3.9
Community or military job	20	-6.4	-6.3	-6.4	PS	-6.3	-6.3	-6.3	-6.3	-6.3	PS	-6.3	PS	-7.9	-6.3	PS	PS
Work from house	21	-4.2	-2.9	-3	-6.9	-3	-3	-3	-3	-3	-5.9	-3	PS	-3.7	-3	-8.2	-3.9
Doing other things	22	-5	-5	-4.8	PS	-5	-5	-5	-5	-5	PS	-5	PS	-5.8	-5	PS	-7.7
Education job	23	-0.7	1.5	0.2	-5.8	0.9	0.9	0.9	0.9	0.9	-5	0.9	-5.9	-2.3	0.9	-6.2	-4.2
Self-employed	24	-1.4	-1.5	-1	-5.8	-1	-1	-1	-1	-1	-4.8	-1	-5.6	-2.6	-1	-6.2	-3.7
Family-owned business	25	-0.8	0.4	0.6	-5.7	1.9	1.9	1.9	1.9	1.9	-4.8	1.9	-5.7	-2.4	1.9	-6.2	-3.8
Unlimited contract	26	-3.7	-3.6	-3.5	PS	-3.6	-3.6	-3.6	-3.6	-3.6	-6.8	-3.6	-7.9	-4.4	-3.6	-7.9	-5.4
Work hours	27	-0.7	0.9	0.9	-5.8	7.8	7.8	7.8	7.8	7.8	-4.9	7.8	-5.8	-2.4	7.8	-6.3	-3.9
Supervising other employees	28	-4.1	-3.2	-3.2	-7.5	-3.3	-3.3	-3.3	-3.3	-3.3	-5.7	-3.3	-8.1	-3.7	-3.3	-8.3	-5.4
General health	29	-0.7	0.9	0.9	-5.8	PC	PC	PC	PC	PC	-4.9	PC	-5.8	-2.4	PC	-6.3	-3.9
Daily activities limitation	30	-4.7	-4.6	-4.5	PS	-4.6	-4.6	-4.6	-4.6	-4.6	PS	-4.6	PS	-4.7	-4.6	PS	-7.2
Left political orientation	31	-4.3	-4.4	-4.3	PS	-4.2	-4.2	-4.2	-4.2	-4.2	-6	-4.2	PS	-4.8	-4.2	PS	-6
Right political orientation	32	-5.7	-5.5	-5.5	PS	-5.6	-5.6	-5.6	-5.6	-5.6	PS	-5.6	PS	-6.2	-5.6	PS	-7.2

Table 3.17: Jointness of Happiness Determinants: Wave 2 (Continued)

Variable	x_h/x_1	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Gender	1	-0.7	-0.7	-0.7	-6.4	-4.2	-5	-0.7	-1.4	-0.8	-3.7	-0.7	-4.1	-0.7	-4.7	-4.3	-5.7
Age	2	0.9	0.9	0.9	-6.3	-2.9	-5	1.5	-1.5	0.4	-3.6	0.9	-3.2	0.9	-4.6	-4.4	-5.5
Country born	3	0.9	0.9	0.9	-6.4	-3	-4.8	0.2	-1	0.6	-3.5	0.9	-3.2	0.9	-4.5	-4.3	-5.5
Minority ethnic group	4	-5.8	-5.8	-5.8	PS	-6.9	PS	-5.8	-5.8	-5.7	PS	-5.8	-7.5	-5.8	PS	PS	PS
Discriminated	5	PC	PC	PC	-6.3	-3	-5	0.9	-1	1.9	-3.6	7.8	-3.3	PC	-4.6	-4.2	-5.6
Income	6	PC	PC	PC	-6.3	-3	-5	0.9	-1	1.9	-3.6	7.8	-3.3	PC	-4.6	-4.2	-5.6
Borrow money	7	PC	PC	PC	-6.3	-3	-5	0.9	-1	1.9	-3.6	7.8	-3.3	PC	-4.6	-4.2	-5.6
Victim of burglary	8	PC	PC	PC	-6.3	-3	-5	0.9	-1	1.9	-3.6	7.8	-3.3	PC	-4.6	-4.2	-5.6
Income feeling	9	PC	PC	PC	-6.3	-3	-5	0.9	-1	1.9	-3.6	7.8	-3.3	PC	-4.6	-4.2	-5.6
Number of household members	10	-4.9	-4.9	-4.9	PS	-5.9	PS	-5	-4.8	-4.8	-6.8	-4.9	-5.7	-4.9	PS	-6	PS
Partner	11	PC	PC	PC	-6.3	-3	-5	0.9	-1	1.9	-3.6	7.8	-3.3	PC	-4.6	-4.2	-5.6
Widowed	12	-5.8	-5.8	-5.8	PS	PS	PS	-5.9	-5.6	-5.7	-7.9	-5.8	-8.1	-5.8	PS	PS	PS
Divorced	13	-2.4	-2.4	-2.4	-7.9	-3.7	-5.8	-2.3	-2.6	-2.4	-4.4	-2.4	-3.7	-2.4	-4.7	-4.8	-6.2
Separated	14	PC	PC	PC	-6.3	-3	-5	0.9	-1	1.9	-3.6	7.8	-3.3	PC	-4.6	-4.2	-5.6
Married	15	-6.3	-6.3	-6.3	PS	-8.2	PS	-6.2	-6.2	-6.2	-7.9	-6.3	-8.3	-6.3	PS	PS	PS
Living with children	16	-3.9	-3.9	-3.9	PS	-3.9	-7.7	-4.2	-3.7	-3.8	-5.4	-3.9	-5.4	-3.9	-7.2	-6	-7.2
Unemployed	17	.	PC	PC	-6.3	-3	-5	0.9	-1	1.9	-3.6	7.8	-3.3	PC	-4.6	-4.2	-5.6
Retired	18	PC	.	PC	-6.3	-3	-5	0.9	-1	1.9	-3.6	7.8	-3.3	PC	-4.6	-4.2	-5.6
Paid work	19	PC	PC	.	-6.3	-3	-5	0.9	-1	1.9	-3.6	7.8	-3.3	PC	-4.6	-4.2	-5.6
Community or military job	20	-6.3	-6.3	-6.3	.	-9.3	PS	-6.3	-7.4	-6.4	PS	-6.3	-7.8	-6.3	PS	PS	PS
Work from house	21	-3	-3	-3	-9.3	.	-7.1	-2.9	-3.2	-3	-4.5	-3	-4.1	-3	-4.9	-4.8	-8.2
Doing other things	22	-5	-5	-5	PS	-7.1	.	-5	-5.1	-4.9	-5.7	-5	-4.8	-5	-5.7	-6.3	PS
Education job	23	0.9	0.9	0.9	-6.3	-2.9	-5	.	-1.4	0.5	-3.8	0.9	-3.4	0.9	-4.7	-4.4	-5.5
Self-employed	24	-1	-1	-1	-7.4	-3.2	-5.1	-1.4	.	-0.9	-6.2	-1	-3.3	-1	-4.6	-4.3	-5.5
Family-owned business	25	1.9	1.9	1.9	-6.4	-3	-4.9	0.5	-0.9	.	-3.7	1.9	-3.3	1.9	-4.6	-4.2	-5.5
Unlimited contract	26	-3.6	-3.6	-3.6	PS	-4.5	-5.7	-3.8	-6.2	-3.7	.	-3.6	-4.9	-3.6	-5.1	-4.9	-8.3
Work hours	27	7.8	7.8	7.8	-6.3	-3	-5	0.9	-1	1.9	-3.6	.	-3.3	7.8	-4.6	-4.2	-5.6
Supervising other employees	28	-3.3	-3.3	-3.3	-7.8	-4.1	-4.8	-3.4	-3.3	-3.3	-4.9	-3.3	.	-3.3	-6.3	-5.5	PS
General health	29	PC	PC	PC	-6.3	-3	-5	0.9	-1	1.9	-3.6	7.8	-3.3	.	-4.6	-4.2	-5.6
Daily activities limitation	30	-4.6	-4.6	-4.6	PS	-4.9	-5.7	-4.7	-4.6	-4.6	-5.1	-4.6	-6.3	-4.6	.	-7	PS
Left political orientation	31	-4.2	-4.2	-4.2	PS	-4.8	-6.3	-4.4	-4.3	-4.2	-4.9	-4.2	-5.5	-4.2	-7	.	PS
Right political orientation	32	-5.6	-5.6	-5.6	PS	-8.2	PS	-5.5	-5.5	-5.5	-8.3	-5.6	PS	-5.6	PS	PS	.

Table 3.18: Jointness of Happiness Determinants: Wave 3

Variable	$\mathbf{x_h/x_l}$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Gender	1	.	-0.8	-0.9	-6	-0.7	-0.7	-0.7	-0.7	-0.7	-4.8	-0.7	-5.8	-2.3	-0.7	-6.3	-3.7
Age	2	-0.8	.	0.1	-5.8	0.9	0.9	0.9	0.9	0.9	-5.3	0.9	-5.9	-2.2	0.9	-6.3	-4.2
Country born	3	-0.9	0.1	.	-5.8	0.9	0.9	0.9	0.9	0.9	-4.7	0.9	-5.6	-2.4	0.9	-6.1	-3.9
Minority ethnic group	4	-6	-5.8	-5.8	.	-5.9	-5.9	-5.9	-5.9	-5.9	PS	-5.9	PS	-6	-5.9	PS	PS
Discriminated	5	-0.7	0.9	0.9	-5.9	.	PC	PC	PC	PC	-4.6	PC	-5.7	-2.4	PC	-6.2	-3.9
Income	6	-0.7	0.9	0.9	-5.9	PC	.	PC	PC	PC	-4.6	PC	-5.7	-2.4	PC	-6.2	-3.9
Borrow money	7	-0.7	0.9	0.9	-5.9	PC	PC	.	PC	PC	-4.6	PC	-5.7	-2.4	PC	-6.2	-3.9
Victim of burglary	8	-0.7	0.9	0.9	-5.9	PC	PC	PC	.	PC	-4.6	PC	-5.7	-2.4	PC	-6.2	-3.9
Income feeling	9	-0.7	0.9	0.9	-5.9	PC	PC	PC	PC	.	-4.6	PC	-5.7	-2.4	PC	-6.2	-3.9
Number of household members	10	-4.8	-5.3	-4.7	PS	-4.6	-4.6	-4.6	-4.6	-4.6	.	-4.6	PS	-5.8	-4.6	PS	PS
Partner	11	-0.7	0.9	0.9	-5.9	PC	PC	PC	PC	PC	-4.6	.	-5.7	-2.4	PC	-6.2	-3.9
Widowed	12	-5.8	-5.9	-5.6	PS	-5.7	-5.7	-5.7	-5.7	-5.7	PS	-5.7	.	-6.6	-5.7	PS	-8.5
Divorced	13	-2.3	-2.2	-2.4	-6	-2.4	-2.4	-2.4	-2.4	-2.4	-5.8	-2.4	-6.6	.	-2.4	-8.4	-4.3
Separated	14	-0.7	0.9	0.9	-5.9	PC	PC	PC	PC	PC	-4.6	PC	-5.7	-2.4	.	-6.2	-3.9
Married	15	-6.3	-6.3	-6.1	PS	-6.2	-6.2	-6.2	-6.2	-6.2	PS	-6.2	PS	-8.4	-6.2	.	PS
Living with children	16	-3.7	-4.2	-3.9	PS	-3.9	-3.9	-3.9	-3.9	-3.9	PS	-3.9	-8.5	-4.3	-3.9	PS	.
Unemployed	17	-0.7	0.9	0.9	-5.9	PC	PC	PC	PC	PC	-4.6	PC	-5.7	-2.4	PC	-6.2	-3.9
Retired	18	-0.7	0.9	0.9	-5.9	10.9	10.9	10.9	10.9	10.9	-4.6	10.9	-5.7	-2.4	10.9	-6.2	-3.9
Paid work	19	-0.7	0.9	0.9	-5.9	PC	PC	PC	PC	PC	-4.6	PC	-5.7	-2.4	PC	-6.2	-3.9
Community or military job	20	-5.4	-5.8	-5.9	PS	-5.8	-5.8	-5.8	-5.8	-5.8	PS	-5.8	PS	-6.9	-5.8	PS	PS
Work from house	21	-4.2	-3.1	-3.1	-5.6	-3.1	-3.1	-3.1	-3.1	-3.1	-5.8	-3.1	-9.3	-3.8	-3.1	PS	-4.4
Doing other things	22	-5	-4.9	-4.9	PS	-4.9	-4.9	-4.9	-4.9	-4.9	-8.3	-4.9	PS	-5.4	-4.9	PS	PS
Education job	23	-0.6	1.4	0.2	-5.8	0.9	0.9	0.9	0.9	0.9	-4.8	0.9	-5.8	-2.4	0.9	-6.2	-4.1
Self-employed	24	-1.3	-1.5	-1	-6	-1	-1	-1	-1	-1	-4.7	-1	-5.5	-2.7	-1	-6.7	-3.8
Family-owned business	25	-0.8	0.5	0.6	-5.8	1.9	1.9	1.9	1.9	1.9	-4.6	1.9	-5.7	-2.4	1.9	-6	-3.9
Unlimited contract	26	-3.7	-3.5	-3.5	-8.5	-3.5	-3.5	-3.5	-3.5	-3.5	-6.1	-3.5	PS	-4.2	-3.5	-8.1	-5.1
Work hours	27	-0.7	0.9	0.9	-5.9	9.4	9.4	9.4	9.4	9.4	-4.6	9.4	-5.7	-2.4	9.4	-6.2	-3.9
Supervising other employees	28	-4.2	-3.2	-3.3	-9.3	-3.3	-3.3	-3.3	-3.3	-3.3	-5.8	-3.3	-8.4	-3.9	-3.3	-6.2	-4.9
General health	29	-0.7	0.9	0.9	-5.9	PC	PC	PC	PC	PC	-4.6	PC	-5.7	-2.4	PC	-6.2	-3.9
Daily activities limitation	30	-4.5	-4.3	-4.4	PS	-4.4	-4.4	-4.4	-4.4	-4.4	-6.9	-4.4	PS	-4.8	-4.4	PS	-6.8
Left political orientation	31	-4.3	-4.4	-4.3	PS	-4.3	-4.3	-4.3	-4.3	-4.3	-5.5	-4.3	PS	-4.9	-4.3	PS	-5.4
Right political orientation	32	-5.7	-5.4	-5.5	PS	-5.5	-5.5	-5.5	-5.5	-5.5	PS	-5.5	PS	-5.7	-5.5	PS	-7

Table 3.19: Jointness of Happiness Determinants: Wave 3 (Continued)

Variable	$\mathbf{x}_h/\mathbf{x}_l$	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Gender	1	-0.7	-0.7	-0.7	-5.4	-4.2	-5	-0.6	-1.3	-0.8	-3.7	-0.7	-4.2	-0.7	-4.5	-4.3	-5.7
Age	2	0.9	0.9	0.9	-5.8	-3.1	-4.9	1.4	-1.5	0.5	-3.5	0.9	-3.2	0.9	-4.3	-4.4	-5.4
Country born	3	0.9	0.9	0.9	-5.9	-3.1	-4.9	0.2	-1	0.6	-3.5	0.9	-3.3	0.9	-4.4	-4.3	-5.5
Minority ethnic group	4	-5.9	-5.9	-5.9	PS	-5.6	PS	-5.8	-6	-5.8	-8.5	-5.9	-9.3	-5.9	PS	PS	PS
Discriminated	5	PC	10.9	PC	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	9.4	-3.3	PC	-4.4	-4.3	-5.5
Income	6	PC	10.9	PC	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	9.4	-3.3	PC	-4.4	-4.3	-5.5
Borrow money	7	PC	10.9	PC	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	9.4	-3.3	PC	-4.4	-4.3	-5.5
Victim of burglary	8	PC	10.9	PC	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	9.4	-3.3	PC	-4.4	-4.3	-5.5
Income feeling	9	PC	10.9	PC	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	9.4	-3.3	PC	-4.4	-4.3	-5.5
Number of household members	10	-4.6	-4.6	-4.6	PS	-5.8	-8.3	-4.8	-4.7	-4.6	-6.1	-4.6	-5.8	-4.6	-6.9	-5.5	PS
Partner	11	PC	10.9	PC	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	9.4	-3.3	PC	-4.4	-4.3	-5.5
Widowed	12	-5.7	-5.7	-5.7	PS	-9.3	PS	-5.8	-5.5	-5.7	PS	-5.7	-8.4	-5.7	PS	PS	PS
Divorced	13	-2.4	-2.4	-2.4	-6.9	-3.8	-5.4	-2.4	-2.7	-2.4	-4.2	-2.4	-3.9	-2.4	-4.8	-4.9	-5.7
Separated	14	PC	10.9	PC	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	9.4	-3.3	PC	-4.4	-4.3	-5.5
Married	15	-6.2	-6.2	-6.2	PS	PS	PS	-6.2	-6.7	-6	-8.1	-6.2	-6.2	-6.2	PS	PS	PS
Living with children	16	-3.9	-3.9	-3.9	PS	-4.4	PS	-4.1	-3.8	-3.9	-5.1	-3.9	-4.9	-3.9	-6.8	-5.4	-7
Unemployed	17	.	10.9	PC	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	9.4	-3.3	PC	-4.4	-4.3	-5.5
Retired	18	10.9	.	10.9	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	9.2	-3.3	10.9	-4.4	-4.3	-5.5
Paid work	19	PC	10.9	.	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	9.4	-3.3	PC	-4.4	-4.3	-5.5
Community or military job	20	-5.8	-5.8	-5.8	.	-7.2	PS	-5.7	-5.8	-5.8	PS	-5.8	PS	-5.8	PS	PS	PS
Work from house	21	-3.1	-3.1	-3.1	-7.2	.	-7.5	-3	-3.2	-3.1	-4.6	-3.1	-3.9	-3.1	-5	-5.3	-7.5
Doing other things	22	-4.9	-4.9	-4.9	PS	-7.5	.	-4.9	-4.8	-4.9	-6.9	-4.9	-6.5	-4.9	PS	PS	PS
Education job	23	0.9	0.9	0.9	-5.7	-3	-4.9	.	-1.4	0.5	-3.8	0.9	-3.3	0.9	-4.3	-4.4	-5.4
Self-employed	24	-1	-1	-1	-5.8	-3.2	-4.8	-1.4	.	-0.9	-5.8	-1	-3.3	-1	-4.4	-4.4	-5.6
Family-owned business	25	1.9	1.9	1.9	-5.8	-3.1	-4.9	0.5	-0.9	.	-3.6	1.9	-3.3	1.9	-4.4	-4.3	-5.4
Unlimited contract	26	-3.5	-3.5	-3.5	PS	-4.6	-6.9	-3.8	-5.8	-3.6	.	-3.5	-4.3	-3.5	-6.6	-5.4	-7.9
Work hours	27	9.4	9.2	9.4	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	.	-3.3	9.4	-4.4	-4.3	-5.5
Supervising other employees	28	-3.3	-3.3	-3.3	PS	-3.9	-6.5	-3.3	-3.3	-3.3	-4.3	-3.3	.	-3.3	-5.5	-5.6	-6.8
General health	29	PC	10.9	PC	-5.8	-3.1	-4.9	0.9	-1	1.9	-3.5	9.4	-3.3	.	-4.4	-4.3	-5.5
Daily activities limitation	30	-4.4	-4.4	-4.4	PS	-5	PS	-4.3	-4.4	-4.4	-6.6	-4.4	-5.5	-4.4	.	-6.2	-7.2
Left political orientation	31	-4.3	-4.3	-4.3	PS	-5.3	PS	-4.4	-4.4	-4.3	-5.4	-4.3	-5.6	-4.3	-6.2	.	PS
Right political orientation	32	-5.5	-5.5	-5.5	PS	-7.5	PS	-5.4	-5.6	-5.4	-7.9	-5.5	-6.8	-5.5	-7.2	PS	.

Table 3.20: Jointness of Happiness Determinants: Wave 4

Variable	x_h/x_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Gender	1	.	-0.8	-0.9	-5.7	-0.7	-0.7	-0.7	-0.7	-0.7	-5.1	-0.7	-6.7	-2.3	-0.7	-6.4	-3.8
Age	2	-0.8	.	0.1	-5.9	0.8	0.8	0.8	0.8	0.8	-5.4	0.8	-6.3	-2.2	0.8	-6.1	-4.4
Country born	3	-0.9	0.1	.	-5.9	0.9	0.9	0.9	0.9	0.9	-4.8	0.9	-5.8	-2.3	0.9	-6.1	-3.9
Minority ethnic group	4	-5.7	-5.9	-5.9	.	-5.9	-5.9	-5.9	-5.9	-5.9	PS	-5.9	PS	-6.3	-5.9	PS	PS
Discriminated	5	-0.7	0.8	0.9	-5.9	.	PC	PC	PC	PC	-4.9	PC	-5.9	-2.3	PC	-6.2	-4
Income	6	-0.7	0.8	0.9	-5.9	PC	.	PC	PC	PC	-4.9	PC	-5.9	-2.3	PC	-6.2	-4
Borrow money	7	-0.7	0.8	0.9	-5.9	PC	PC	.	PC	PC	-4.9	PC	-5.9	-2.3	PC	-6.2	-4
Victim of burglary	8	-0.7	0.8	0.9	-5.9	PC	PC	PC	.	PC	-4.9	PC	-5.9	-2.3	PC	-6.2	-4
Income feeling	9	-0.7	0.8	0.9	-5.9	PC	PC	PC	PC	.	-4.9	PC	-5.9	-2.3	PC	-6.2	-4
Number of household members	10	-5.1	-5.4	-4.8	PS	-4.9	-4.9	-4.9	-4.9	-4.9	.	-4.9	PS	-5.3	-4.9	PS	PS
Partner	11	-0.7	0.8	0.9	-5.9	PC	PC	PC	PC	PC	-4.9	.	-5.9	-2.3	PC	-6.2	-4
Widowed	12	-6.7	-6.3	-5.8	PS	-5.9	-5.9	-5.9	-5.9	-5.9	PS	-5.9	.	-7.2	-5.9	PS	-5.9
Divorced	13	-2.3	-2.2	-2.3	-6.3	-2.3	-2.3	-2.3	-2.3	-2.3	-5.3	-2.3	-7.2	.	-2.3	-7.1	-4.9
Separated	14	-0.7	0.8	0.9	-5.9	PC	PC	PC	PC	PC	-4.9	PC	-5.9	-2.3	.	-6.2	-4
Married	15	-6.4	-6.1	-6.1	PS	-6.2	-6.2	-6.2	-6.2	-6.2	PS	-6.2	PS	-7.1	-6.2	.	PS
Living with children	16	-3.8	-4.4	-3.9	PS	-4	-4	-4	-4	-4	PS	-4	-5.9	-4.9	-4	PS	.
Unemployed	17	-0.7	0.8	0.9	-5.9	PC	PC	PC	PC	PC	-4.9	PC	-5.9	-2.3	PC	-6.2	-4
Retired	18	-0.7	0.8	0.9	-5.9	PC	PC	PC	PC	PC	-4.9	PC	-5.9	-2.3	PC	-6.2	-4
Paid work	19	-0.7	0.8	0.9	-5.9	PC	PC	PC	PC	PC	-4.9	PC	-5.9	-2.3	PC	-6.2	-4
Community or military job	20	-6.2	-5.8	-5.7	PS	-5.8	-5.8	-5.8	-5.8	-5.8	PS	-5.8	PS	-6.3	-5.8	PS	PS
Work from house	21	-4.1	-3	-3	-10	-3	-3	-3	-3	-3	-5.4	-3	-8	-3.6	-3	-8	-4.4
Doing other things	22	-5.1	-4.7	-4.8	PS	-4.8	-4.8	-4.8	-4.8	-4.8	PS	-4.8	PS	-5.2	-4.8	PS	-5.6
Education job	23	-0.6	1.4	0.1	-5.9	0.9	0.9	0.9	0.9	0.9	-5	0.9	-6.1	-2.3	0.9	-6	-4.4
Self-employed	24	-1.4	-1.5	-1	-5.9	-1	-1	-1	-1	-1	-4.8	-1	-5.8	-2.6	-1	-6.6	-3.8
Family-owned business	25	-0.8	0.4	0.6	-5.8	1.9	1.9	1.9	1.9	1.9	-4.9	1.9	-5.9	-2.4	1.9	-6.1	-3.9
Unlimited contract	26	-3.7	-3.6	-3.5	PS	-3.6	-3.6	-3.6	-3.6	-3.6	-6.5	-3.6	-7.7	-4.3	-3.6	-8.8	-4.7
Work hours	27	-0.7	0.8	0.9	-5.9	8.1	8.1	8.1	8.1	8.1	-4.9	8.1	-5.9	-2.3	8.1	-6.2	-4
Supervising other employees	28	-4.1	-3.3	-3.2	-9.4	-3.3	-3.3	-3.3	-3.3	-3.3	-6.4	-3.3	-8	-4	-3.3	PS	-4.7
General health	29	-0.7	0.8	0.9	-5.9	PC	PC	PC	PC	PC	-4.9	PC	-5.9	-2.3	PC	-6.2	-4
Daily activities limitation	30	-4.6	-4.4	-4.3	PS	-4.4	-4.4	-4.4	-4.4	-4.4	PS	-4.4	PS	-4.6	-4.4	PS	-7.8
Left political orientation	31	-4.3	-4.3	-4.3	PS	-4.3	-4.3	-4.3	-4.3	-4.3	-7.1	-4.3	-7.6	-4.8	-4.3	PS	-5.8
Right political orientation	32	-5.2	-5.1	-5.2	PS	-5.2	-5.2	-5.2	-5.2	-5.2	PS	-5.2	PS	-5.5	-5.2	PS	-8.1

Table 3.21: Jointness of Happiness Determinants: Wave 4 (Continued)

Variable	x_h/x_1	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Gender	1	-0.7	-0.7	-0.7	-6.2	-4.1	-5.1	-0.6	-1.4	-0.8	-3.7	-0.7	-4.1	-0.7	-4.6	-4.3	-5.2
Age	2	0.8	0.8	0.8	-5.8	-3	-4.7	1.4	-1.5	0.4	-3.6	0.8	-3.3	0.8	-4.4	-4.3	-5.1
Country born	3	0.9	0.9	0.9	-5.7	-3	-4.8	0.1	-1	0.6	-3.5	0.9	-3.2	0.9	-4.3	-4.3	-5.2
Minority ethnic group	4	-5.9	-5.9	-5.9	PS	-10	PS	-5.9	-5.9	-5.8	PS	-5.9	-9.4	-5.9	PS	PS	PS
Discriminated	5	PC	PC	PC	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	8.1	-3.3	PC	-4.4	-4.3	-5.2
Income	6	PC	PC	PC	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	8.1	-3.3	PC	-4.4	-4.3	-5.2
Borrow money	7	PC	PC	PC	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	8.1	-3.3	PC	-4.4	-4.3	-5.2
Victim of burglary	8	PC	PC	PC	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	8.1	-3.3	PC	-4.4	-4.3	-5.2
Income feeling	9	PC	PC	PC	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	8.1	-3.3	PC	-4.4	-4.3	-5.2
Number of household members	10	-4.9	-4.9	-4.9	PS	-5.4	PS	-5	-4.8	-4.9	-6.5	-4.9	-6.4	-4.9	PS	-7.1	PS
Partner	11	PC	PC	PC	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	8.1	-3.3	PC	-4.4	-4.3	-5.2
Widowed	12	-5.9	-5.9	-5.9	PS	-8	PS	-6.1	-5.8	-5.9	-7.7	-5.9	-8	-5.9	PS	-7.6	PS
Divorced	13	-2.3	-2.3	-2.3	-6.3	-3.6	-5.2	-2.3	-2.6	-2.4	-4.3	-2.3	-4	-2.3	-4.6	-4.8	-5.5
Separated	14	PC	PC	PC	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	8.1	-3.3	PC	-4.4	-4.3	-5.2
Married	15	-6.2	-6.2	-6.2	PS	-8	PS	-6	-6.6	-6.1	-8.8	-6.2	PS	-6.2	PS	PS	PS
Living with children	16	-4	-4	-4	PS	-4.4	-5.6	-4.4	-3.8	-3.9	-4.7	-4	-4.7	-4	-7.8	-5.8	-8.1
Unemployed	17	.	PC	PC	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	8.1	-3.3	PC	-4.4	-4.3	-5.2
Retired	18	PC	.	PC	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	8.1	-3.3	PC	-4.4	-4.3	-5.2
Paid work	19	PC	PC	.	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	8.1	-3.3	PC	-4.4	-4.3	-5.2
Community or military job	20	-5.8	-5.8	-5.8	.	-6.9	PS	-5.8	-6.5	-5.7	PS	-5.8	-5.9	-5.8	PS	PS	PS
Work from house	21	-3	-3	-3	-6.9	.	-6	-2.9	-3.1	-3	-4.3	-3	-4.2	-3	-5.1	-5	-6.9
Doing other things	22	-4.8	-4.8	-4.8	PS	-6	.	-4.7	-5.1	-4.8	-6.8	-4.8	-5.7	-4.8	PS	PS	PS
Education job	23	0.9	0.9	0.9	-5.8	-2.9	-4.7	.	-1.4	0.5	-3.8	0.9	-3.4	0.9	-4.4	-4.3	-5.1
Self-employed	24	-1	-1	-1	-6.5	-3.1	-5.1	-1.4	.	-0.9	-6.2	-1	-3.1	-1	-4.5	-4.5	-5.2
Family-owned business	25	1.9	1.9	1.9	-5.7	-3	-4.8	0.5	-0.9	.	-3.6	1.9	-3.2	1.9	-4.4	-4.3	-5.1
Unlimited contract	26	-3.6	-3.6	-3.6	PS	-4.3	-6.8	-3.8	-6.2	-3.6	.	-3.6	-4.8	-3.6	-5.4	-5.1	-7.8
Work hours	27	8.1	8.1	8.1	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	.	-3.3	8.1	-4.4	-4.3	-5.2
Supervising other employees	28	-3.3	-3.3	-3.3	-5.9	-4.2	-5.7	-3.4	-3.1	-3.2	-4.8	-3.3	.	-3.3	-5.3	-6.1	-5.3
General health	29	PC	PC	PC	-5.8	-3	-4.8	0.9	-1	1.9	-3.6	8.1	-3.3	.	-4.4	-4.3	-5.2
Daily activities limitation	30	-4.4	-4.4	-4.4	PS	-5.1	PS	-4.4	-4.5	-4.4	-5.4	-4.4	-5.3	-4.4	.	-5.6	PS
Left political orientation	31	-4.3	-4.3	-4.3	PS	-5	PS	-4.3	-4.5	-4.3	-5.1	-4.3	-6.1	-4.3	-5.6	.	-5.3
Right political orientation	32	-5.2	-5.2	-5.2	PS	-6.9	PS	-5.1	-5.2	-5.1	-7.8	-5.2	-5.3	-5.2	PS	-5.3	.

Table 3.22: Jointness of Happiness Determinants: Wave 1-3

Variable	x_h/x_1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Gender	1	.	-3.1	-2.3	-4.3	3.8	3.8	3.8	3.8	3.8	-4.4	3.8	-5.9	-3.4	3.8	-5.6	-3.1
Age	2	-3.1	.	-4	-4.7	-3.1	-3.1	-3.1	-3.1	-3.1	-6	-3.1	-7.9	-3.9	-3.1	-8.5	-4.9
Country born	3	-2.3	-4	.	-3.3	-2.3	-2.3	-2.3	-2.3	-2.3	-5.1	-2.3	-5.8	-3.7	-2.3	-6.5	-3.6
Minority ethnic group	4	-4.3	-4.7	-3.3	.	-4.3	-4.3	-4.3	-4.3	-4.3	-7.1	-4.3	PS	-5.6	-4.3	-7.5	-5.1
Discriminated	5	3.8	-3.1	-2.3	-4.3	.	PC	PC	6.8	PC	-4.4	PC	-5.9	-3.4	PC	-5.6	-3.1
Income	6	3.8	-3.1	-2.3	-4.3	PC	.	PC	6.8	PC	-4.4	PC	-5.9	-3.4	PC	-5.6	-3.1
Borrow money	7	3.8	-3.1	-2.3	-4.3	PC	PC	.	6.8	PC	-4.4	PC	-5.9	-3.4	PC	-5.6	-3.1
Victim of burglary	8	3.8	-3.1	-2.3	-4.3	6.8	6.8	6.8	.	6.8	-4.4	6.8	-5.9	-3.4	6.8	-5.6	-3.1
Income feeling	9	3.8	-3.1	-2.3	-4.3	PC	PC	PC	6.8	.	-4.4	PC	-5.9	-3.4	PC	-5.6	-3.1
Number of household members	10	-4.4	-6	-5.1	-7.1	-4.4	-4.4	-4.4	-4.4	-4.4	.	-4.4	PS	-5.3	-4.4	PS	-6.4
Partner	11	3.8	-3.1	-2.3	-4.3	PC	PC	PC	6.8	PC	-4.4	.	-5.9	-3.4	PC	-5.6	-3.1
Widowed	12	-5.9	-7.9	-5.8	PS	-5.9	-5.9	-5.9	-5.9	-5.9	PS	-5.9	.	-7.5	-5.9	PS	-6.1
Divorced	13	-3.4	-3.9	-3.7	-5.6	-3.4	-3.4	-3.4	-3.4	-3.4	-5.3	-3.4	-7.5	.	-3.4	-7.5	-4.7
Separated	14	3.8	-3.1	-2.3	-4.3	PC	PC	PC	6.8	PC	-4.4	PC	-5.9	-3.4	.	-5.6	-3.1
Married	15	-5.6	-8.5	-6.5	-7.5	-5.6	-5.6	-5.6	-5.6	-5.6	PS	-5.6	PS	-7.5	-5.6	.	-6.8
Living with children	16	-3.1	-4.9	-3.6	-5.1	-3.1	-3.1	-3.1	-3.1	-3.1	-6.4	-3.1	-6.1	-4.7	-3.1	-6.8	.
Unemployed	17	3.8	-3.1	-2.3	-4.3	PC	PC	PC	6.8	PC	-4.4	PC	-5.9	-3.4	PC	-5.6	-3.1
Retired	18	3.8	-3.1	-2.3	-4.3	PC	PC	PC	6.8	PC	-4.4	PC	-5.9	-3.4	PC	-5.6	-3.1
Paid work	19	3.8	-3.1	-2.3	-4.3	PC	PC	PC	6.8	PC	-4.4	PC	-5.9	-3.4	PC	-5.6	-3.1
Community or military job	20	-5.5	-6.2	-6.8	PS	-5.5	-5.5	-5.5	-5.5	-5.5	-7	-5.5	PS	-6.8	-5.5	PS	-6.9
Work from house	21	-3.7	-4.6	-4	-5.5	-3.6	-3.6	-3.6	-3.6	-3.6	-5	-3.6	-5.6	-4.9	-3.6	PS	-4
Doing other things	22	-5.8	-8	-6.9	-6.8	-5.8	-5.8	-5.8	-5.8	-5.8	PS	-5.8	PS	-7.4	-5.8	PS	-6.5
Education job	23	-1	-2.2	-2.4	-4.3	-1	-1	-1	-1	-1	-4.4	-1	-6.4	-4.1	-1	-5.6	-3.7
Self-employed	24	-4.7	-5.1	-5.1	-6	-4.7	-4.7	-4.7	-4.7	-4.7	PS	-4.7	PS	-5.2	-4.7	PS	-6
Family-owned business	25	0.8	-3.2	-2.3	-4.3	0.9	0.9	0.9	0.9	0.9	-4.4	0.9	-5.8	-3.4	0.9	-5.5	-3.2
Unlimited contract	26	-4.8	-6.2	-5.4	-6.1	-4.8	-4.8	-4.8	-4.8	-4.8	-6.9	-4.8	PS	-5.6	-4.8	PS	-5
Work hours	27	1.7	-3.2	-2.4	-4.4	1.8	1.8	1.8	1.8	1.8	-4.4	1.8	-5.9	-3.4	1.8	-5.7	-3.1
Supervising other employees	28	-4	-4.3	-4.4	-6.7	-4	-4	-4	-4	-4	-5.4	-4	PS	-5.2	-4	PS	-4.8
General health	29	3.8	-3.1	-2.3	-4.3	PC	PC	PC	6.8	PC	-4.4	PC	-5.9	-3.4	PC	-5.6	-3.1
Daily activities limitation	30	-2.9	-3.6	-3.8	-5.1	-2.9	-2.9	-2.9	-2.9	-2.9	-5	-2.9	-9.4	-4.5	-2.9	-6.3	-3.8
Left political orientation	31	-4.3	-4.5	-4.7	-6.2	-4.3	-4.3	-4.3	-4.3	-4.3	-8.4	-4.3	PS	-5.3	-4.3	-7	-5.6
Right political orientation	32	-2.5	-3.8	-3.3	-5.2	-2.5	-2.5	-2.5	-2.5	-2.5	-5.1	-2.5	-9.2	-3.9	-2.5	-5.7	-3.9

Table 3.23: Jointness of Happiness Determinants: Wave 1-3 (Continued)

Variable	x_h/x_1	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Gender	1	3.8	3.8	3.8	-5.5	-3.7	-5.8	-1	-4.7	0.8	-4.8	1.7	-4	3.8	-2.9	-4.3	-2.5
Age	2	-3.1	-3.1	-3.1	-6.2	-4.6	-8	-2.2	-5.1	-3.2	-6.2	-3.2	-4.3	-3.1	-3.6	-4.5	-3.8
Country born	3	-2.3	-2.3	-2.3	-6.8	-4	-6.9	-2.4	-5.1	-2.3	-5.4	-2.4	-4.4	-2.3	-3.8	-4.7	-3.3
Minority ethnic group	4	-4.3	-4.3	-4.3	PS	-5.5	-6.8	-4.3	-6	-4.3	-6.1	-4.4	-6.7	-4.3	-5.1	-6.2	-5.2
Discriminated	5	PC	PC	PC	-5.5	-3.6	-5.8	-1	-4.7	0.9	-4.8	1.8	-4	PC	-2.9	-4.3	-2.5
Income	6	PC	PC	PC	-5.5	-3.6	-5.8	-1	-4.7	0.9	-4.8	1.8	-4	PC	-2.9	-4.3	-2.5
Borrow money	7	PC	PC	PC	-5.5	-3.6	-5.8	-1	-4.7	0.9	-4.8	1.8	-4	PC	-2.9	-4.3	-2.5
Victim of burglary	8	6.8	6.8	6.8	-5.5	-3.6	-5.8	-1	-4.7	0.9	-4.8	1.8	-4	6.8	-2.9	-4.3	-2.5
Income feeling	9	PC	PC	PC	-5.5	-3.6	-5.8	-1	-4.7	0.9	-4.8	1.8	-4	PC	-2.9	-4.3	-2.5
Number of household members	10	-4.4	-4.4	-4.4	-7	-5	PS	-4.4	PS	-4.4	-6.9	-4.4	-5.4	-4.4	-5	-8.4	-5.1
Partner	11	PC	PC	PC	-5.5	-3.6	-5.8	-1	-4.7	0.9	-4.8	1.8	-4	PC	-2.9	-4.3	-2.5
Widowed	12	-5.9	-5.9	-5.9	PS	-5.6	PS	-6.4	PS	-5.8	PS	-5.9	PS	-5.9	-9.4	PS	-9.2
Divorced	13	-3.4	-3.4	-3.4	-6.8	-4.9	-7.4	-4.1	-5.2	-3.4	-5.6	-3.4	-5.2	-3.4	-4.5	-5.3	-3.9
Separated	14	PC	PC	PC	-5.5	-3.6	-5.8	-1	-4.7	0.9	-4.8	1.8	-4	PC	-2.9	-4.3	-2.5
Married	15	-5.6	-5.6	-5.6	PS	PS	PS	-5.6	PS	-5.5	PS	-5.7	PS	-5.6	-6.3	-7	-5.7
Living with children	16	-3.1	-3.1	-3.1	-6.9	-4	-6.5	-3.7	-6	-3.2	-5	-3.1	-4.8	-3.1	-3.8	-5.6	-3.9
Unemployed	17	.	PC	PC	-5.5	-3.6	-5.8	-1	-4.7	0.9	-4.8	1.8	-4	PC	-2.9	-4.3	-2.5
Retired	18	PC	.	PC	-5.5	-3.6	-5.8	-1	-4.7	0.9	-4.8	1.8	-4	PC	-2.9	-4.3	-2.5
Paid work	19	PC	PC	.	-5.5	-3.6	-5.8	-1	-4.7	0.9	-4.8	1.8	-4	PC	-2.9	-4.3	-2.5
Community or military job	20	-5.5	-5.5	-5.5	.	-7.7	PS	-5.5	PS	-5.4	PS	-5.5	PS	-5.5	-5.8	-6.6	-6.8
Work from house	21	-3.6	-3.6	-3.6	-7.7	.	PS	-3.4	-5.4	-3.5	-5.9	-3.5	-5.3	-3.6	-4.5	-5.7	-4.3
Doing other things	22	-5.8	-5.8	-5.8	PS	PS	.	-6.1	PS	-5.7	PS	-5.8	PS	-5.8	-6.7	PS	-6
Education job	23	-1	-1	-1	-5.5	-3.4	-6.1	.	-5	-1.2	-5.5	-1.4	-4.1	-1	-3.2	-4.5	-2.6
Self-employed	24	-4.7	-4.7	-4.7	PS	-5.4	PS	-5	.	-4.6	-7.7	-4.6	-8.3	-4.7	-5.3	-6.1	-4.8
Family-owned business	25	0.9	0.9	0.9	-5.4	-3.5	-5.7	-1.2	-4.6	.	-4.8	0.7	-4	0.9	-2.8	-4.3	-2.5
Unlimited contract	26	-4.8	-4.8	-4.8	PS	-5.9	PS	-5.5	-7.7	-4.8	.	-4.6	-7.4	-4.8	-5.1	-6	-4.9
Work hours	27	1.8	1.8	1.8	-5.5	-3.5	-5.8	-1.4	-4.6	0.7	-4.6	.	-4.2	1.8	-2.9	-4.3	-2.5
Supervising other employees	28	-4	-4	-4	PS	-5.3	PS	-4.1	-8.3	-4	-7.4	-4.2	.	-4	-4.8	-6.4	-5.2
General health	29	PC	PC	PC	-5.5	-3.6	-5.8	-1	-4.7	0.9	-4.8	1.8	-4	.	-2.9	-4.3	-2.5
Daily activities limitation	30	-2.9	-2.9	-2.9	-5.8	-4.5	-6.7	-3.2	-5.3	-2.8	-5.1	-2.9	-4.8	-2.9	.	-4.3	-3.6
Left political orientation	31	-4.3	-4.3	-4.3	-6.6	-5.7	PS	-4.5	-6.1	-4.3	-6	-4.3	-6.4	-4.3	-4.3	.	-5.3
Right political orientation	32	-2.5	-2.5	-2.5	-6.8	-4.3	-6	-2.6	-4.8	-2.5	-4.9	-2.5	-5.2	-2.5	-3.6	-5.3	.

Table 3.24: Jointness of Happiness Determinants: Wave 1-4

Variable	x_h/x_l	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Gender	1	.	-0.8	-0.9	-5.7	-0.6	-0.6	-0.6	-0.6	-0.6	-5.1	-0.6	-5.9	-2.3	-0.6	-6.6	-3.8
Age	2	-0.8	.	0.1	-5.7	0.9	0.9	0.9	0.9	0.9	-5.3	0.9	-5.9	-2.2	0.9	-6.3	-4.2
Country born	3	-0.9	0.1	.	-5.7	0.9	0.9	0.9	0.9	0.9	-4.6	0.9	-5.7	-2.3	0.9	-6.4	-3.8
Minority ethnic group	4	-5.7	-5.7	-5.7	.	-5.7	-5.7	-5.7	-5.7	-5.7	PS	-5.7	PS	-7	-5.7	PS	PS
Discriminated	5	-0.6	0.9	0.9	-5.7	.	PC	PC	PC	PC	-4.7	PC	-5.7	-2.3	PC	-6.3	-3.9
Income	6	-0.6	0.9	0.9	-5.7	PC	.	PC	PC	PC	-4.7	PC	-5.7	-2.3	PC	-6.3	-3.9
Borrow money	7	-0.6	0.9	0.9	-5.7	PC	PC	.	PC	PC	-4.7	PC	-5.7	-2.3	PC	-6.3	-3.9
Victim of burglary	8	-0.6	0.9	0.9	-5.7	PC	PC	PC	.	PC	-4.7	PC	-5.7	-2.3	PC	-6.3	-3.9
Income feeling	9	-0.6	0.9	0.9	-5.7	PC	PC	PC	PC	.	-4.7	PC	-5.7	-2.3	PC	-6.3	-3.9
Number of household members	10	-5.1	-5.3	-4.6	PS	-4.7	-4.7	-4.7	-4.7	-4.7	.	-4.7	PS	-5.1	-4.7	PS	PS
Partner	11	-0.6	0.9	0.9	-5.7	PC	PC	PC	PC	PC	-4.7	.	-5.7	-2.3	PC	-6.3	-3.9
Widowed	12	-5.9	-5.9	-5.7	PS	-5.7	-5.7	-5.7	-5.7	-5.7	PS	-5.7	.	-6.7	-5.7	PS	-5.7
Divorced	13	-2.3	-2.2	-2.3	-7	-2.3	-2.3	-2.3	-2.3	-2.3	-5.1	-2.3	-6.7	.	-2.3	-7.2	-5.4
Separated	14	-0.6	0.9	0.9	-5.7	PC	PC	PC	PC	PC	-4.7	PC	-5.7	-2.3	.	-6.3	-3.9
Married	15	-6.6	-6.3	-6.4	PS	-6.3	-6.3	-6.3	-6.3	-6.3	PS	-6.3	PS	-7.2	-6.3	.	PS
Living with children	16	-3.8	-4.2	-3.8	PS	-3.9	-3.9	-3.9	-3.9	-3.9	PS	-3.9	-5.7	-5.4	-3.9	PS	.
Unemployed	17	-0.6	0.9	0.9	-5.7	PC	PC	PC	PC	PC	-4.7	PC	-5.7	-2.3	PC	-6.3	-3.9
Retired	18	-0.6	0.9	0.9	-5.7	PC	PC	PC	PC	PC	-4.7	PC	-5.7	-2.3	PC	-6.3	-3.9
Paid work	19	-0.6	0.9	0.9	-5.7	PC	PC	PC	PC	PC	-4.7	PC	-5.7	-2.3	PC	-6.3	-3.9
Community or military job	20	-6.1	-6.1	-5.9	PS	-6	-6	-6	-6	-6	PS	-6	PS	-6	-6	PS	PS
Work from house	21	-4.1	-3.1	-3.1	-6.4	-3.1	-3.1	-3.1	-3.1	-3.1	-5.1	-3.1	-7	-3.5	-3.1	PS	-4
Doing other things	22	-5	-4.8	-4.9	PS	-4.9	-4.9	-4.9	-4.9	-4.9	-8	-4.9	PS	-6	-4.9	PS	-6.2
Education job	23	-0.6	1.4	0.1	-5.7	0.9	0.9	0.9	0.9	0.9	-5	0.9	-5.7	-2.3	0.9	-6.2	-4.2
Self-employed	24	-1.3	-1.5	-0.9	-5.8	-0.9	-0.9	-0.9	-0.9	-0.9	-4.5	-0.9	-5.9	-2.7	-0.9	-6.3	-3.7
Family-owned business	25	-0.8	0.5	0.6	-5.6	1.9	1.9	1.9	1.9	1.9	-4.6	1.9	-5.7	-2.4	1.9	-6.2	-3.8
Unlimited contract	26	-3.8	-3.6	-3.6	PS	-3.7	-3.7	-3.7	-3.7	-3.7	-6	-3.7	-10.1	-4.3	-3.7	PS	-5.7
Work hours	27	-0.6	0.9	0.9	-5.7	8.2	8.2	8.2	8.2	8.2	-4.7	8.2	-5.7	-2.3	8.2	-6.3	-3.9
Supervising other employees	28	-4	-3.3	-3.5	-7.8	-3.5	-3.5	-3.5	-3.5	-3.5	-5.3	-3.5	PS	-4	-3.5	PS	-5.2
General health	29	-0.6	0.9	0.9	-5.7	PC	PC	PC	PC	PC	-4.7	PC	-5.7	-2.3	PC	-6.3	-3.9
Daily activities limitation	30	-4.5	-4.2	-4.3	PS	-4.3	-4.3	-4.3	-4.3	-4.3	-9.9	-4.3	PS	-4.8	-4.3	PS	-5.5
Left political orientation	31	-4.4	-4.5	-4.3	PS	-4.4	-4.4	-4.4	-4.4	-4.4	-6.8	-4.4	-6.1	-5	-4.4	PS	-6.3
Right political orientation	32	-5.7	-5.6	-5.5	PS	-5.6	-5.6	-5.6	-5.6	-5.6	PS	-5.6	PS	-6.7	-5.6	PS	-6

Table 3.25: Jointness of Happiness Determinants: Wave 1-4 (Continued)

Variable	x_h/x_l	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Gender	1	-0.6	-0.6	-0.6	-6.1	-4.1	-5	-0.6	-1.3	-0.8	-3.8	-0.6	-4	-0.6	-4.5	-4.4	-5.7
Age	2	0.9	0.9	0.9	-6.1	-3.1	-4.8	1.4	-1.5	0.5	-3.6	0.9	-3.3	0.9	-4.2	-4.5	-5.6
Country born	3	0.9	0.9	0.9	-5.9	-3.1	-4.9	0.1	-0.9	0.6	-3.6	0.9	-3.5	0.9	-4.3	-4.3	-5.5
Minority ethnic group	4	-5.7	-5.7	-5.7	PS	-6.4	PS	-5.7	-5.8	-5.6	PS	-5.7	-7.8	-5.7	PS	PS	PS
Discriminated	5	PC	PC	PC	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	8.2	-3.5	PC	-4.3	-4.4	-5.6
Income	6	PC	PC	PC	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	8.2	-3.5	PC	-4.3	-4.4	-5.6
Borrow money	7	PC	PC	PC	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	8.2	-3.5	PC	-4.3	-4.4	-5.6
Victim of burglary	8	PC	PC	PC	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	8.2	-3.5	PC	-4.3	-4.4	-5.6
Income feeling	9	PC	PC	PC	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	8.2	-3.5	PC	-4.3	-4.4	-5.6
Number of household members	10	-4.7	-4.7	-4.7	PS	-5.1	-8	-5	-4.5	-4.6	-6	-4.7	-5.3	-4.7	-9.9	-6.8	PS
Partner	11	PC	PC	PC	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	8.2	-3.5	PC	-4.3	-4.4	-5.6
Widowed	12	-5.7	-5.7	-5.7	PS	-7	PS	-5.7	-5.9	-5.7	-10.1	-5.7	PS	-5.7	PS	-6.1	PS
Divorced	13	-2.3	-2.3	-2.3	-6	-3.5	-6	-2.3	-2.7	-2.4	-4.3	-2.3	-4	-2.3	-4.8	-5	-6.7
Separated	14	PC	PC	PC	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	8.2	-3.5	PC	-4.3	-4.4	-5.6
Married	15	-6.3	-6.3	-6.3	PS	PS	PS	-6.2	-6.3	-6.2	PS	-6.3	PS	-6.3	PS	PS	PS
Living with children	16	-3.9	-3.9	-3.9	PS	-4	-6.2	-4.2	-3.7	-3.8	-5.7	-3.9	-5.2	-3.9	-5.5	-6.3	-6
Unemployed	17	.	PC	PC	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	8.2	-3.5	PC	-4.3	-4.4	-5.6
Retired	18	PC	.	PC	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	8.2	-3.5	PC	-4.3	-4.4	-5.6
Paid work	19	PC	PC	.	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	8.2	-3.5	PC	-4.3	-4.4	-5.6
Community or military job	20	-6	-6	-6	.	-8.1	PS	-6	-5.8	-5.9	-7.3	-6	PS	-6	PS	PS	PS
Work from house	21	-3.1	-3.1	-3.1	-8.1	.	-6.1	-3	-3.3	-3.1	-4.3	-3.1	-4.1	-3.1	-5.3	-5.1	-7
Doing other things	22	-4.9	-4.9	-4.9	PS	-6.1	.	-4.9	-5	-4.8	-8.9	-4.9	-5.3	-4.9	-5.7	-6.4	PS
Education job	23	0.9	0.9	0.9	-6	-3	-4.9	.	-1.4	0.5	-3.8	0.9	-3.4	0.9	-4.2	-4.4	-5.6
Self-employed	24	-0.9	-0.9	-0.9	-5.8	-3.3	-5	-1.4	.	-0.9	-5.9	-0.9	-3.5	-0.9	-4.5	-4.4	-5.8
Family-owned business	25	1.9	1.9	1.9	-5.9	-3.1	-4.8	0.5	-0.9	.	-3.7	1.9	-3.4	1.9	-4.3	-4.3	-5.6
Unlimited contract	26	-3.7	-3.7	-3.7	-7.3	-4.3	-8.9	-3.8	-5.9	-3.7	.	-3.7	-4.9	-3.7	-5.8	-6.8	PS
Work hours	27	8.2	8.2	8.2	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	.	-3.5	8.2	-4.3	-4.4	-5.6
Supervising other employees	28	-3.5	-3.5	-3.5	PS	-4.1	-5.3	-3.4	-3.5	-3.4	-4.9	-3.5	.	-3.5	-4.5	-5.7	-9.7
General health	29	PC	PC	PC	-6	-3.1	-4.9	0.9	-0.9	1.9	-3.7	8.2	-3.5	.	-4.3	-4.4	-5.6
Daily activities limitation	30	-4.3	-4.3	-4.3	PS	-5.3	-5.7	-4.2	-4.5	-4.3	-5.8	-4.3	-4.5	-4.3	.	-7.1	PS
Left political orientation	31	-4.4	-4.4	-4.4	PS	-5.1	-6.4	-4.4	-4.4	-4.3	-6.8	-4.4	-5.7	-4.4	-7.1	.	PS
Right political orientation	32	-5.6	-5.6	-5.6	PS	-7	PS	-5.6	-5.8	-5.6	PS	-5.6	-9.7	-5.6	PS	PS	.

3.7 BMA Analysis Across Countries

In this section, I run the BMA analysis for each country using all waves combined and compare PIP measures across different countries. My aim is to investigate the differences in happiness determinants across different countries if there is any. Tables 3.26 through 3.53 present the results for the selected countries. My findings show that different countries look at life satisfaction differently. A factor that makes the people of country A happy does not necessarily makes the people of country B happy with their life. For instance, Spanish people think that age is not an important factor determining how happy they are while German think of age as a key factor of happiness.

3.8 Conclusions

This paper investigates the effective happiness determinants for European countries. I employ Bayesian model averaging—a practical and theoretically well-established method—to estimate the posterior coefficients as well as posterior inclusion probabilities. The value of estimated coefficients across different waves of dataset and also across individual countries is illustrated in the previous sections, using BMA paradigm and incorporating recent Markov Chain Monte Carlo (MCMC) samplers to search over the model space. Overall, income, being employed, general health, being socially active, having a partner, not being discriminated, and being involved in volunteering activities are among the key determinants of happiness.

Table 3.26: Bayesian Model Averaging Analysis: Austria

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Gender	1.0000	-0.4744	0.0887	Voted last election	0.0608	0.0105	0.0517
General health	1.0000	-0.5430	0.0554	Supervising other employees	0.0553	0.0063	0.0340
Socializing	1.0000	0.1621	0.0308	Paid work	0.0524	-0.0133	0.0753
Borrow money	0.9996	0.2001	0.0398	Country born	0.0470	0.0097	0.0586
Income feeling	0.9995	-0.3236	0.0643	Living with children	0.0434	0.0040	0.0340
Retired	0.9794	0.6384	0.1614	Separated	0.0392	-0.0152	0.1192
Discriminated	0.9399	-0.6887	0.2628	Victim of burglary	0.0337	0.0047	0.0388
Number of household members	0.7422	0.0778	0.0544	Married	0.0335	0.0002	0.0262
Left political orientation	0.7210	-0.2668	0.1847	Divorced	0.0318	0.0019	0.0306
Religious	0.6715	0.0447	0.0356	Right political orientation	0.0300	-0.0005	0.0269
Widowed	0.4346	-0.2016	0.2578	Age	0.0292	0.0000	0.0008
Education level (2)	0.4229	0.1226	0.1602	Pray	0.0290	-0.0003	0.0052
Attending religious services	0.3630	-0.0336	0.0491	Social activities	0.0282	0.0010	0.0101
Member of a political party	0.3392	0.0953	0.1484	Daily activities limitation	0.0281	-0.0013	0.0213
Politics scale	0.3047	0.0278	0.0453	Minority ethnic group	0.0258	0.0033	0.0405
Income	0.1510	-0.0073	0.0198	Intimate	0.0257	0.0029	0.0316
Work from house	0.1027	0.0208	0.0717	Interested in politics	0.0252	-0.0005	0.0092
Partner	0.0807	0.0141	0.0589	Family-owned business	0.0248	-0.0059	0.0709
Unemployed	0.0752	-0.0340	0.1486	Work hours	0.0243	-0.0001	0.0007
Self-employed	0.0681	-0.0163	0.0721	Education level (1)	0.0211	-0.0003	0.0666
Unlimited job contract	0.0672	0.0099	0.0457	Education job	0.0209	-0.0014	0.0367

Table 3.27: Bayesian Model Averaging Analysis: Belgium

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.5710	0.0506	Right political orientation	0.0932	0.0159	0.0624
General health	1.0000	-0.4067	0.0522	Number of household members	0.0896	-0.0050	0.0190
Intimate	0.9994	0.5403	0.1165	Widowed	0.0879	-0.0292	0.1131
Religious	0.9946	0.0836	0.0199	Gender	0.0872	-0.0111	0.0425
Country born	0.9609	0.5048	0.1801	Left political orientation	0.0548	0.0070	0.0469
Politics scale	0.8881	0.0654	0.0307	Member of a political party	0.0538	0.0099	0.0523
Discriminated	0.7739	-0.3749	0.2431	Divorced	0.0526	-0.0116	0.0638
Married	0.6166	0.2125	0.1850	Income	0.0509	-0.0015	0.0083
Education level (1)	0.5826	0.2004	0.1930	Interested in politics	0.0442	0.0025	0.0152
Retired	0.5658	0.1834	0.1802	Supervising other employees	0.0439	-0.0040	0.0246
Social activities	0.4961	0.0515	0.0581	Attending religious services	0.0399	-0.0014	0.0089
Socializing	0.4506	0.0335	0.0414	Daily activities limitation	0.0383	-0.0041	0.0288
Living with children	0.4420	-0.1075	0.1350	Pray	0.0380	0.0007	0.0052
Partner	0.3809	0.1304	0.1800	Education level (2)	0.0374	0.0036	0.0280
Voted last election	0.3783	-0.1298	0.1855	Unlimited job contract	0.0351	-0.0036	0.0245
Minority ethnic group	0.2575	-0.1378	0.2619	Education job	0.0338	0.0079	0.0585
Borrow money	0.2421	0.0168	0.0333	Separated	0.0255	0.0065	0.0690
Paid work	0.1680	-0.0391	0.0978	Unemployed	0.0254	-0.0029	0.0358
Work from house	0.1397	0.0250	0.0705	Self-employed	0.0218	-0.0001	0.0176
Age	0.1194	0.0008	0.0026	Family-owned business	0.0177	0.0009	0.0413
Victim of burglary	0.1037	-0.0160	0.0543	Work hours	0.0167	0.0000	0.0003

Table 3.28: Bayesian Model Averaging Analysis: Bulgaria

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	0.9993	-0.7208	0.1693	Income	0.0714	0.0106	0.0599
Borrow money	0.8316	0.2740	0.1573	Divorced	0.0692	0.0303	0.1571
Discriminated	0.7698	-0.9887	0.6624	Supervising other employees	0.0688	0.0143	0.0824
Education level (2)	0.7079	-0.5554	0.4259	Left political orientation	0.0563	-0.0093	0.0646
Daily activities limitation	0.6716	-0.5351	0.4385	Social activities	0.0553	0.0046	0.0304
Victim of burglary	0.5154	-0.3235	0.3605	Family-owned business	0.0549	0.0315	0.2323
Married	0.5087	0.4042	0.4430	Gender	0.0516	-0.0044	0.0519
Partner	0.4877	0.3906	0.4501	Work hours	0.0513	-0.0003	0.0028
Number of household members	0.4874	-0.0943	0.1115	Intimate	0.0508	0.0123	0.0872
General health	0.3366	-0.1083	0.1739	Age	0.0495	0.0003	0.0025
Minority ethnic group	0.1590	-0.1020	0.2834	Unemployed	0.0487	0.0037	0.0951
Paid work	0.1470	-0.0615	0.1755	Widowed	0.0485	0.0049	0.0946
Separated	0.1294	-0.1696	0.5347	Politics scale	0.0480	0.0007	0.0100
Voted last election	0.1110	0.0392	0.1372	Right political orientation	0.0478	0.0046	0.0573
Socializing	0.1090	0.0095	0.0337	Member of a political party	0.0476	-0.0061	0.0813
Attending religious services	0.0897	-0.0105	0.0435	Work from house	0.0471	0.0041	0.0565
Retired	0.0837	0.0207	0.1011	Pray	0.0465	-0.0014	0.0138
Country born	0.0832	-0.0910	0.3931	Interested in politics	0.0425	-0.0008	0.0245
Unlimited job contract	0.0786	-0.0224	0.1020	Education level (1)	0.0417	-0.0086	0.1343
Self-employed	0.0785	0.0499	0.2393	Education job	0.0405	-0.0134	0.2220
Living with children	0.0728	-0.0131	0.0938	Religious	0.0377	-0.0005	0.0084

Table 3.29: Bayesian Model Averaging Analysis: Switzerland

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Discriminated	1.0000	-0.8013	0.1333	Income	0.0754	-0.0017	0.0067
Income feeling	1.0000	-0.5509	0.0398	Minority ethnic group	0.0740	-0.0151	0.0634
General health	1.0000	-0.5477	0.0384	Education level (2)	0.0646	0.0080	0.0362
Religious	1.0000	0.0601	0.0112	Living with children	0.0520	-0.0050	0.0260
Socializing	1.0000	0.1485	0.0229	Self-employed	0.0407	-0.0046	0.0279
Unemployed	0.9997	-0.9018	0.2099	Interested in politics	0.0342	-0.0015	0.0102
Partner	0.9977	0.4486	0.0671	Daily activities limitation	0.0338	-0.0034	0.0230
Country born	0.9566	0.2579	0.0882	Married	0.0315	0.0038	0.0303
Retired	0.8565	0.4453	0.2235	Work hours	0.0304	-0.0001	0.0004
Victim of burglary	0.8293	-0.1931	0.1080	Education job	0.0302	0.0032	0.0325
Work from house	0.5764	0.1091	0.1049	Pray	0.0288	0.0004	0.0032
Paid work	0.4992	-0.1387	0.1568	Number of household members	0.0244	-0.0006	0.0056
Age	0.3031	0.0021	0.0035	Supervising other employees	0.0233	-0.0007	0.0095
Intimate	0.2339	0.0756	0.1517	Divorced	0.0230	-0.0017	0.0183
Social activities	0.2026	0.0144	0.0319	Widowed	0.0226	0.0022	0.0226
Gender	0.1977	-0.0254	0.0572	Attending religious services	0.0219	-0.0003	0.0039
Education level (1)	0.1891	-0.0615	0.1421	Right political orientation	0.0211	0.0005	0.0109
Separated	0.1724	-0.0643	0.1575	Unlimited job contract	0.0204	0.0011	0.0119
Borrow money	0.1210	0.0057	0.0174	Family-owned business	0.0198	0.0004	0.0248
Left political orientation	0.1049	-0.0115	0.0383	Voted last election	0.0186	0.0006	0.0099
Politics scale	0.0803	0.0020	0.0080	Member of a political party	0.0172	-0.0003	0.0117

Table 3.30: Bayesian Model Averaging Analysis: Cyprus

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Divorced	0.9993	-2.2034	0.4696	Daily activities limitation	0.1407	-0.0716	0.2161
General health	0.9748	-0.4518	0.1474	Member of a political party	0.1306	0.0372	0.1182
Religious	0.9432	0.1497	0.0563	Retired	0.1151	0.0388	0.1446
Work hours	0.9285	-0.0236	0.0098	Intimate	0.1145	0.0486	0.1734
Socializing	0.9030	0.1524	0.0706	Interested in politics	0.1028	-0.0116	0.0448
Income feeling	0.6388	-0.2048	0.1810	Partner	0.0896	0.0177	0.1005
Age	0.6168	0.0112	0.0104	Education job	0.0843	-0.0570	0.2539
Borrow money	0.3609	0.0609	0.0924	Right political orientation	0.0785	0.0099	0.0891
Separated	0.3003	0.9218	1.6268	Married	0.0777	0.0142	0.0921
Widowed	0.2893	-0.2405	0.4382	Attending religious services	0.0724	-0.0063	0.0337
Pray	0.2694	-0.0266	0.0511	Number of household members	0.0707	0.0037	0.0218
Left political orientation	0.2571	0.0892	0.1769	Family-owned business	0.0681	0.0258	0.1431
Gender	0.2430	0.0885	0.1844	Discriminated	0.0669	-0.0296	0.1800
Politics scale	0.2066	-0.0107	0.0263	Unlimited job contract	0.0660	-0.0093	0.0615
Education level (1)	0.1874	-0.0837	0.2088	Living with children	0.0635	0.0055	0.0488
Voted last election	0.1871	-0.1114	0.2760	Income	0.0554	0.0013	0.0144
Victim of burglary	0.1857	0.0940	0.2320	Paid work	0.0534	0.0006	0.0564
Self-employed	0.1850	0.0740	0.1863	Minority ethnic group	0.0529	0.0131	0.1555
Work from house	0.1847	0.0646	0.1638	Country born	0.0519	-0.0007	0.0844
Unemployed	0.1693	-0.1116	0.2981	Social activities	0.0519	0.0009	0.0244
Supervising other employees	0.1506	0.0446	0.1291	Education level (2)	0.0458	0.0012	0.0608

Table 3.31: Bayesian Model Averaging Analysis: Czech Republic

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.5427	0.1077	Partner	0.0794	0.0144	0.0788
General health	1.0000	-0.5863	0.1055	Left political orientation	0.0791	-0.0223	0.0988
Right political orientation	0.9451	0.6192	0.2119	Politics scale	0.0790	0.0050	0.0256
Voted last election	0.9155	0.5012	0.2173	Unemployed	0.0723	-0.0356	0.1697
Social activities	0.8592	0.1987	0.1048	Widowed	0.0691	-0.0084	0.0861
Religious	0.7540	0.0636	0.0439	Daily activities limitation	0.0653	-0.0130	0.0685
Separated	0.7511	-1.3148	0.9134	Self-employed	0.0629	0.0218	0.1196
Discriminated	0.7077	-0.6757	0.5165	Number of household members	0.0593	-0.0041	0.0263
Divorced	0.6491	-0.4351	0.3731	Pray	0.0576	-0.0011	0.0126
Victim of burglary	0.6240	-0.3303	0.2981	Attending religious services	0.0531	-0.0010	0.0166
Age	0.5939	0.0089	0.0085	Paid work	0.0523	-0.0036	0.0509
Interested in politics	0.4152	-0.1008	0.1363	Work hours	0.0479	0.0002	0.0017
Borrow money	0.3175	0.0509	0.0849	Country born	0.0474	0.0097	0.0829
Work from house	0.2915	0.1195	0.2116	Living with children	0.0466	-0.0039	0.0423
Gender	0.2362	-0.0714	0.1471	Education level (1)	0.0459	-0.0193	0.2157
Unlimited job contract	0.1837	0.0589	0.1442	Education job	0.0446	0.0157	0.1647
Member of a political party	0.1644	0.0911	0.2420	Income	0.0425	-0.0014	0.0128
Married	0.1621	0.0532	0.1438	Education level (2)	0.0422	-0.0028	0.0465
Socializing	0.1219	0.0094	0.0304	Supervising other employees	0.0410	-0.0039	0.0414
Retired	0.1002	0.0251	0.1007	Minority ethnic group	0.0384	-0.0081	0.0866
Intimate	0.0871	0.0217	0.0889	Family-owned business	0.0380	-0.0063	0.1143

Table 3.32: Bayesian Model Averaging Analysis: Germany

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.7903	0.0491	Victim of burglary	0.0593	-0.0096	0.0451
Unemployed	1.0000	-0.6839	0.1420	Left political orientation	0.0475	-0.0055	0.0336
General health	1.0000	-0.6433	0.0372	Voted last election	0.0390	0.0044	0.0280
Socializing	1.0000	0.1390	0.0241	Married	0.0359	0.0068	0.0467
Age	0.9986	0.0139	0.0036	Self-employed	0.0302	-0.0032	0.0256
Borrow money	0.9928	0.1175	0.0297	Right political orientation	0.0298	0.0015	0.0241
Partner	0.9821	0.3883	0.0919	Daily activities limitation	0.0277	-0.0025	0.0195
Social activities	0.9733	0.1358	0.0416	Supervising other employees	0.0268	-0.0016	0.0148
Discriminated	0.9109	-0.5262	0.2252	Work from house	0.0260	0.0014	0.0137
Politics scale	0.9020	0.0559	0.0245	Unlimited job contract	0.026	0.0018	0.0158
Retired	0.8256	0.2729	0.1529	Living with children	0.0244	-0.0015	0.0192
Work hours	0.7726	-0.0052	0.0034	Education level (1)	0.0228	-0.0047	0.0530
Religious	0.6820	0.0285	0.0223	Minority ethnic group	0.0215	-0.0028	0.0295
Number of household members	0.2092	0.0144	0.0313	Interested in politics	0.0206	-0.0003	0.0059
Income	0.1486	0.0049	0.0131	Pray	0.0197	-0.0002	0.0029
Education job	0.1172	0.0424	0.1305	Intimate	0.0194	0.0004	0.0222
Attending religious services	0.0993	-0.0050	0.0172	Education level (2)	0.0188	-0.0010	0.0161
Country born	0.0934	-0.0189	0.0674	Gender	0.0183	-0.0001	0.0095
Paid work	0.0765	-0.0136	0.0534	Divorced	0.0145	-0.0006	0.0151
Separated	0.0676	-0.0283	0.1224	Member of a political party	0.0135	-0.0007	0.0172
Family-owned business	0.0632	0.0303	0.1374	Widowed	0.0117	0.0001	0.0139

Table 3.33: Bayesian Model Averaging Analysis: Denmark

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.3011	0.0468	Left political orientation	0.0508	-0.0034	0.0230
General health	1.0000	-0.3652	0.0366	Right political orientation	0.0499	0.0020	0.0215
Socializing	0.9999	0.1142	0.0222	Gender	0.0452	-0.0029	0.0174
Age	0.9974	0.0115	0.0022	Pray	0.0379	0.0005	0.0035
Borrow money	0.9946	0.1048	0.0257	Paid work	0.0368	-0.0024	0.0211
Partner	0.9897	0.3112	0.0713	Widowed	0.0351	0.0045	0.0329
Discriminated	0.9862	-0.5774	0.1607	Family-owned business	0.0330	0.0096	0.0770
Intimate	0.889	0.3210	0.1502	Interested in politics	0.0326	0.0012	0.0093
Social activities	0.6933	0.0650	0.0501	Unlimited job contract	0.0304	-0.0017	0.0140
Unemployed	0.6085	-0.2593	0.2367	Income	0.0298	-0.0004	0.0032
Country born	0.3310	-0.1118	0.1759	Supervising other employees	0.0287	-0.0014	0.0121
Politics scale	0.2871	0.0094	0.0165	Married	0.0287	0.0028	0.0243
Daily activities limitation	0.2822	-0.0479	0.0850	Victim of burglary	0.0283	0.0011	0.0116
Education level (2)	0.2150	0.0348	0.0740	Living with children	0.0229	0.0003	0.0107
Minority ethnic group	0.1303	0.0551	0.1602	Self-employed	0.0209	0.0000	0.0134
Retired	0.1124	0.0188	0.0607	Attending religious services	0.0204	0.0001	0.0034
Separated	0.1011	-0.0422	0.1454	Voted last election	0.0199	0.0004	0.0176
Work hours	0.0760	-0.0003	0.0010	Member of a political party	0.0196	0.0007	0.0139
Education job	0.0737	0.0147	0.0615	Work from house	0.0185	0.0003	0.0081
Number of household members	0.0653	0.0028	0.0130	Religious	0.0182	0.0000	0.0016
Education level (1)	0.0597	0.0214	0.1032	Divorced	0.0170	0.0005	0.0136

Table 3.34: Bayesian Model Averaging Analysis: Spain

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.4694	0.0512	Education job	0.0343	0.0075	0.0568
General health	1.0000	-0.3774	0.0473	Income	0.0289	0.0004	0.0041
Religious	1.0000	0.0900	0.0162	Daily activities limitation	0.0283	-0.0029	0.0269
Socializing	1.0000	0.1497	0.0267	Living with children	0.0248	0.0008	0.0152
Intimate	1.0000	0.7986	0.1465	Paid work	0.0245	-0.0011	0.0207
Married	0.9022	0.4569	0.1730	Member of a political party	0.0244	-0.0021	0.0353
Social activities	0.7179	0.0974	0.0716	Voted last election	0.0239	0.0007	0.0158
Unemployed	0.6645	-0.3451	0.2835	Attending religious services	0.0234	0.0002	0.0046
Discriminated	0.3600	-0.1745	0.2593	Work hours	0.0226	0.0000	0.0005
Unlimited job contract	0.2899	0.0562	0.0984	Education level (1)	0.0218	-0.0010	0.0150
Number of household members	0.1556	0.0096	0.0253	Minority ethnic group	0.0216	-0.0005	0.0351
Partner	0.1248	0.0529	0.1532	Separated	0.0216	-0.0001	0.0406
Interested in politics	0.0980	0.0081	0.0283	Pray	0.0215	-0.0002	0.0031
Age	0.0842	-0.0004	0.0016	Divorced	0.0212	-0.0012	0.0351
Education level (2)	0.0753	0.0107	0.0456	Retired	0.0209	0.0007	0.0197
Politics scale	0.0749	0.0023	0.0097	Work from house	0.0208	0.0005	0.0157
Left political orientation	0.0635	-0.0075	0.0350	Widowed	0.0208	0.0019	0.0296
Family-owned business	0.0586	-0.0280	0.1366	Right political orientation	0.0202	0.0012	0.0157
Borrow money	0.0439	0.0018	0.0109	Country born	0.0193	-0.0003	0.0202
Gender	0.0373	-0.0033	0.0223	Victim of burglary	0.0179	-0.0001	0.0117
Self-employed	0.0352	0.0034	0.0338	Supervising other employees	0.0176	0.0000	0.0118

Table 3.35: Bayesian Model Averaging Analysis: Finland

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.5136	0.0411	Living with children	0.0640	0.0063	0.0281
General health	1.0000	-0.4335	0.0439	Right political orientation	0.0611	-0.0079	0.0397
Religious	1.0000	0.0879	0.0116	Interested in politics	0.0591	0.0028	0.0138
Retired	0.9973	0.3729	0.0682	Family-owned business	0.0581	0.0120	0.0585
Discriminated	0.9907	-0.3837	0.1007	Supervising other employees	0.0532	-0.0045	0.0233
Socializing	0.9855	0.0770	0.0212	Number of household members	0.0480	0.0016	0.0086
Partner	0.9671	0.3411	0.0925	Education job	0.0453	0.0054	0.0317
Intimate	0.8974	0.2876	0.1301	Divorced	0.0416	-0.0048	0.0289
Politics scale	0.7111	0.0448	0.0323	Education level (2)	0.0403	0.0037	0.0232
Daily activities limitation	0.5820	-0.1064	0.1017	Widowed	0.0361	0.0045	0.0311
Social activities	0.4123	0.0285	0.0378	Paid work	0.0290	-0.0019	0.0200
Left political orientation	0.3667	-0.0994	0.1406	Minority ethnic group	0.0240	-0.0036	0.0444
Separated	0.3551	-0.2671	0.3985	Pray	0.0226	0.0001	0.0022
Borrow money	0.2976	0.0155	0.0264	Age	0.0201	0.0000	0.0004
Gender	0.158	-0.0174	0.0451	Member of a political party	0.0192	0.0011	0.0140
Married	0.1301	0.0241	0.0722	Country born	0.0186	0.0023	0.0294
Unemployed	0.1244	-0.0319	0.0956	Education level (1)	0.0171	0.0005	0.0102
Income	0.1089	-0.0031	0.0102	Voted last election	0.0165	0.0002	0.0089
Victim of burglary	0.1003	-0.0102	0.0348	Self-employed	0.0165	0.0002	0.0100
Work hours	0.0736	0.0002	0.0009	Unlimited job contract	0.0161	-0.0002	0.0069
Attending religious services	0.0669	0.0028	0.0124	Work from house	0.0152	-0.0004	0.0165

Table 3.36: Bayesian Model Averaging Analysis: France

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-1.0408	0.0898	Paid work	0.0532	0.0063	0.0543
General health	1.0000	-0.4973	0.0716	Income	0.0511	0.0014	0.0086
Social activities	0.9996	0.3629	0.0760	Number of household members	0.0472	0.0026	0.0161
Discriminated	0.9902	-0.7584	0.1977	Victim of burglary	0.0464	-0.0066	0.0397
Socializing	0.9464	0.1459	0.0543	Country born	0.0385	0.0052	0.0470
Religious	0.9266	0.0772	0.0328	Unlimited job contract	0.0379	-0.0044	0.0322
Retired	0.8968	0.6017	0.2967	Member of a political party	0.0379	-0.0093	0.0804
Voted last election	0.7112	-0.2953	0.2219	Attending religious services	0.0371	-0.0012	0.0121
Age	0.6952	-0.0127	0.0098	Interested in politics	0.0358	-0.0018	0.0154
Left political orientation	0.5556	-0.2193	0.2177	Work hours	0.0353	0.0001	0.0008
Borrow money	0.4949	0.0622	0.0709	Education level (2)	0.0346	0.0035	0.0370
Right political orientation	0.3324	0.1147	0.1794	Minority ethnic group	0.0319	0.0061	0.0599
Widowed	0.2915	-0.1297	0.2276	Family-owned business	0.0319	0.0152	0.1273
Politics scale	0.1715	0.0120	0.0300	Living with children	0.0314	0.0009	0.0259
Pray	0.1599	0.0091	0.0239	Education job	0.0307	-0.0068	0.0760
Gender	0.1428	0.0295	0.0830	Separated	0.0301	-0.0085	0.0950
Education level (1)	0.1109	-0.0304	0.1004	Intimate	0.0293	-0.0035	0.0360
Unemployed	0.0903	-0.0383	0.1447	Divorced	0.0288	-0.0022	0.0365
Self-employed	0.0767	0.0206	0.0875	Married	0.0261	-0.0006	0.0254
Daily activities limitation	0.0640	-0.0112	0.0556	Supervising other employees	0.0257	-0.0002	0.018
Partner	0.0601	0.0098	0.0511	Work from house	0.0237	0.0025	0.0389

Table 3.37: Bayesian Model Averaging Analysis: Great Britain

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.6356	0.0413	Supervising other employess	0.0689	-0.0070	0.0300
Retired	1.0000	0.4476	0.0911	Divorced	0.0639	-0.0096	0.0436
General health	1.0000	-0.4069	0.0327	Education job	0.0614	0.0168	0.0767
Socializing	1.0000	0.1211	0.0202	Family-owned business	0.0472	0.0154	0.0859
Social activities	1.0000	0.1740	0.0314	Self-employed	0.0391	0.0048	0.0300
Unemployed	0.9998	-0.7638	0.1555	Pray	0.0378	0.0008	0.0049
Religious	0.9997	0.0563	0.0106	Gender	0.0318	0.0023	0.0168
Partner	0.9992	0.3907	0.0862	Living with children	0.0314	0.0025	0.0221
Victim of burglary	0.9917	-0.2691	0.0690	Daily activities limitation	0.0314	-0.0032	0.0224
Intimate	0.9663	0.3959	0.1271	Income	0.0287	-0.0004	0.0035
Separated	0.9237	-0.5305	0.2156	Paid work	0.0263	0.0024	0.0208
Politics scale	0.8834	0.0564	0.0256	Left political orientation	0.0261	0.0008	0.0199
Education level (1)	0.8771	0.2238	0.1076	Country born	0.0207	0.0002	0.0148
Borrow money	0.8012	0.0621	0.0372	Education level (2)	0.0205	0.0002	0.0112
Number of household members	0.6565	0.0573	0.0473	Work from house	0.0202	-0.0013	0.0143
Discriminated	0.4566	-0.1059	0.1282	Voted last election	0.0202	0.0002	0.0096
Age	0.1176	-0.0007	0.0021	Widowed	0.0180	-0.0005	0.0156
Right political orientation	0.1154	0.0215	0.0662	Unlimited job contract	0.0149	-0.0002	0.0074
Minority ethnic group	0.1069	-0.0254	0.0825	Married	0.0136	0.0003	0.0123
Work hours	0.0839	-0.0003	0.0012	Member of a political party	0.0132	0.0009	0.0194
Interested in politics	0.0693	0.0042	0.0179	Attending religious services	0.0130	-0.0001	0.0029

Table 3.38: Bayesian Model Averaging Analysis: Greece

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.5336	0.0634	Separated	0.0866	-0.0681	0.2570
General health	1.0000	-0.4958	0.0606	Education job	0.0777	0.0461	0.1849
Religious	0.9969	0.0984	0.0233	Work hours	0.0502	-0.0002	0.0011
Politics scale	0.9866	0.1008	0.0246	Income	0.0494	-0.0016	0.0091
Social activities	0.9109	0.2056	0.0891	Attending religious services	0.0439	0.0024	0.015
Borrow money	0.8970	0.1276	0.0581	Age	0.0402	0.0001	0.0011
Intimate	0.8829	0.5192	0.2491	Family-owned business	0.0339	-0.0020	0.0537
Married	0.8204	0.3564	0.2033	Self-employed	0.0327	0.0017	0.0214
Minority ethnic group	0.6275	-0.4887	0.4273	Daily activities limitation	0.0292	0.0033	0.0320
Socializing	0.5809	0.0584	0.0571	Living with children	0.0286	0.0024	0.0262
Unemployed	0.3106	-0.1600	0.2678	Left political orientation	0.0283	-0.0047	0.0421
Divorced	0.2869	-0.1577	0.2776	Supervising other employees	0.0279	-0.0010	0.0193
Retired	0.2602	0.0761	0.1432	Unlimited job contract	0.0278	0.0014	0.0180
Discriminated	0.2545	-0.1361	0.2582	Education level (2)	0.0249	-0.0025	0.0261
Voted last election	0.2486	0.0921	0.1784	Country born	0.0230	-0.0016	0.0310
Partner	0.1818	0.0539	0.1590	Right political orientation	0.0229	0.0013	0.0331
Victim of burglary	0.1640	-0.0398	0.1012	Number of household members	0.0226	0.0002	0.0066
Paid work	0.1526	-0.0392	0.1058	Work from house	0.0226	0.0024	0.0288
Gender	0.1375	-0.0265	0.0753	Pray	0.0217	-0.0002	0.0046
Education level (1)	0.0936	-0.0201	0.0733	Member of a political party	0.0213	0.0017	0.0264
Widowed	0.0905	0.0321	0.1196	Interested in politics	0.0167	0.0002	0.0066

Table 3.39: Bayesian Model Averaging Analysis: Hungary

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.9164	0.1596	Married	0.0794	-0.0014	0.1180
General health	0.9570	-0.4612	0.1728	Work hours	0.0791	-0.0009	0.0045
Social activities	0.9272	0.4210	0.1720	Member of a political party	0.0763	0.0688	0.3557
Unemployed	0.9011	-1.2191	0.5671	Number of household members	0.0741	0.0053	0.0313
Divorced	0.6141	-0.5478	0.5091	Attending religious services	0.0638	-0.0036	0.0263
Intimate	0.6020	0.8540	0.8099	Right political orientation	0.0631	0.0052	0.1035
Education job	0.5426	0.8262	0.8758	Self-employed	0.0616	0.0283	0.1759
Borrow money	0.2656	0.0605	0.1160	Income	0.0615	0.0026	0.0175
Widowed	0.2391	0.1613	0.3354	Family-owned business	0.0608	-0.0765	0.5028
Education level (2)	0.2214	-0.0971	0.2120	Minority ethnic group	0.0582	-0.0224	0.1483
Politics scale	0.2166	-0.0195	0.0445	Interested in politics	0.0579	0.0067	0.0390
Paid work	0.1846	-0.0893	0.2223	Country born	0.0568	-0.0252	0.1901
Retired	0.1460	0.0641	0.1907	Unlimited job contract	0.0549	0.0061	0.0667
Socializing	0.1425	0.0185	0.0560	Pray	0.0500	-0.0010	0.0121
Left political orientation	0.1242	0.0496	0.1663	Supervising other employees	0.0500	-0.0089	0.0751
Education level (1)	0.1183	0.0770	0.2645	Gender	0.0494	0.0048	0.0547
Discriminated	0.0968	-0.0613	0.2355	Living with children	0.0486	-0.0053	0.0595
Victim of burglary	0.0953	-0.0388	0.156	Daily activities limitation	0.0470	-0.0030	0.0728
Partner	0.0890	-0.0195	0.1215	Work from house	0.0440	-0.0017	0.0495
Religious	0.0872	0.0044	0.0188	Voted last election	0.0433	-0.0043	0.0715
Age	0.0827	0.0007	0.0036	Separated	0.0366	-0.0055	0.2599

Table 3.40: Bayesian Model Averaging Analysis: Ireland

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.6019	0.0608	Socializing	0.1339	0.0071	0.0206
General health	1.0000	-0.4956	0.0487	Number of household members	0.1193	-0.0069	0.0213
Religious	0.997	0.0821	0.0185	Left political orientation	0.0953	0.0194	0.0701
Discriminated	0.9967	-0.8483	0.1977	Education job	0.0812	0.0291	0.1143
Social activities	0.9948	0.2016	0.0455	Unlimited job contract	0.0802	-0.0100	0.0395
Intimate	0.9842	0.5447	0.1533	Widowed	0.0530	0.0096	0.0560
Politics scale	0.965	0.0858	0.0294	Daily activities limitation	0.0483	0.0067	0.0382
Unemployed	0.8643	-0.5299	0.2769	Right political orientation	0.0427	0.0033	0.0365
Retired	0.7400	0.3096	0.209	Member of a political party	0.0333	0.0057	0.0417
Borrow money	0.6582	0.0643	0.0533	Gender	0.0294	0.0016	0.0171
Partner	0.6083	0.1838	0.1658	Income	0.0293	-0.0003	0.0037
Married	0.331	0.0918	0.1427	Country born	0.0278	0.0027	0.0257
Separated	0.2958	-0.1352	0.2327	Attending religious services	0.0273	-0.0008	0.0072
Paid work	0.2912	-0.0825	0.1416	Family-owned business	0.0251	0.0036	0.0448
Education level (1)	0.2736	0.0742	0.1348	Victim of burglary	0.0249	-0.0014	0.0168
Education level (2)	0.2680	0.0645	0.1193	Minority ethnic group	0.0239	-0.0022	0.0348
Living with children	0.2113	-0.0414	0.0893	Voted last election	0.0206	0.0000	0.0145
Interested in politics	0.1864	0.0170	0.0398	Self-employed	0.0182	0.0007	0.0158
Supervising other employees	0.1856	-0.0310	0.0729	Divorced	0.0181	0.0018	0.0414
Work from house	0.1415	0.0267	0.0743	Work hours	0.0177	0.0000	0.0004
Age	0.1410	0.0011	0.0031	Pray	0.0174	0.0000	0.0030

Table 3.41: Bayesian Model Averaging Analysis: Israel

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Discriminated	1.0000	-0.9063	0.1658	Right political orientation	0.0392	0.0022	0.0526
Income feeling	1.0000	-0.5941	0.0631	Separated	0.0356	0.0104	0.1187
General health	1.0000	-0.3609	0.0627	Interested in politics	0.0355	-0.0007	0.0116
Intimate	1.0000	0.7161	0.1587	Self-employed	0.0352	-0.0034	0.0355
Religious	0.9877	0.0808	0.0212	Social activities	0.0351	0.0021	0.0159
Politics scale	0.9838	0.0960	0.0310	Living with children	0.0348	0.0033	0.0303
Socializing	0.9613	0.1418	0.0484	Divorced	0.0348	-0.0046	0.0484
Partner	0.2740	0.0819	0.1494	Family-owned business	0.0313	-0.0132	0.1323
Education level (1)	0.2442	0.2485	0.4908	Unemployed	0.0309	-0.0054	0.0463
Minority ethnic group	0.2381	-0.0843	0.1699	Unlimited job contract	0.0301	0.0000	0.0188
Married	0.2319	0.0653	0.1347	Income	0.0291	0.0001	0.0049
Education level (2)	0.2239	-0.0738	0.1547	Attending religious services	0.0287	-0.0001	0.0073
Gender	0.1707	0.0385	0.0967	Victim of burglary	0.0283	0.0013	0.0261
Country born	0.1437	0.0349	0.0976	Paid work	0.0277	-0.0011	0.0218
Left political orientation	0.1325	0.0479	0.1426	Retired	0.0276	0.0021	0.0305
Daily activities limitation	0.0752	0.0192	0.0826	Education job	0.0274	-0.0022	0.0325
Widowed	0.0571	-0.0193	0.1007	Work from house	0.0271	-0.0019	0.0290
Borrow money	0.0459	0.0023	0.0151	Supervising other employees	0.0267	0.0015	0.0214
Number of household members	0.0458	0.0016	0.0100	Work hours	0.0263	0.0000	0.0006
Pray	0.0417	-0.0010	0.0094	Member of a political party	0.0241	0.0014	0.0314
Age	0.0413	0.0002	0.0013	Voted last election	0.0210	0.0008	0.0216

Table 3.42: Bayesian Model Averaging Analysis: Italy

Variable	PIP	Post Mean	Post SD	Variable	PIP	Post Mean	Post SD
(i)	(ii)	(iii)	(iv)	(v)	(vi)		
Income feeling	1.0000	-0.6384	0.1170	Minority ethnic group	0.0642	0.0477	0.2590
Socializing	1.0000	0.2533	0.0456	Pray	0.0640	0.0026	0.0133
Religious	1.0000	0.1734	0.0338	Education level (1)	0.0634	-0.0133	0.0745
General health	0.9783	-0.3838	0.1149	Right political orientation	0.0615	0.0101	0.0543
Income	0.6769	0.0713	0.0578	Daily activities limitation	0.0562	0.0126	0.0745
Borrow money	0.6743	0.1327	0.1082	Attending religious services	0.0556	-0.0032	0.0180
Partner	0.6033	0.4947	0.4301	Unlimited job contract	0.0536	0.0064	0.0431
Married	0.4345	0.3441	0.4190	Gender	0.0524	-0.0068	0.0439
Age	0.2499	-0.0030	0.0061	Victim of burglary	0.0505	0.0059	0.0450
Country born	0.1919	-0.2094	0.4945	Interested in politics	0.0504	0.0034	0.0232
Unemployed	0.1658	-0.0941	0.2477	Self-employed	0.0477	-0.0056	0.0446
Separated	0.1464	-0.1197	0.3402	Supervising other employees	0.0474	0.0061	0.0434
Intimate	0.1304	-0.0383	0.1171	Member of a political party	0.0460	-0.0118	0.0870
Work from house	0.1179	-0.0533	0.1760	Widowed	0.0437	-0.0053	0.0730
Number of household members	0.0976	0.0098	0.0369	Social activities	0.0429	0.0023	0.0174
Work hours	0.0933	0.0006	0.0024	Living with children	0.0426	-0.0051	0.0457
Left political orientation	0.0801	-0.0149	0.0649	Family-owned business	0.0410	0.0034	0.1494
Politics scale	0.0797	0.0034	0.0144	Education job	0.0400	0.0002	0.0938
Retired	0.0785	0.0234	0.1237	Voted last election	0.0394	-0.0065	0.0608
Divorced	0.0752	-0.0383	0.1760	Discriminated	0.0371	-0.0069	0.1003
Paid work	0.0713	0.014	0.0729	Education level (2)	0.0361	-0.0029	0.0335

Table 3.43: Bayesian Model Averaging Analysis: Luxemburg

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.8692	0.1005	Politics scale	0.0582	0.0017	0.0123
General health	1.0000	-0.5556	0.0771	Pray	0.0546	0.0017	0.0102
Unemployed	0.8798	-1.4181	0.7025	Family-owned business	0.0542	-0.0362	0.2265
Socializing	0.8397	0.1106	0.0615	Right political orientation	0.0535	-0.0057	0.0479
Partner	0.6348	0.3312	0.2854	Living with children	0.0530	-0.0056	0.0427
Education level (1)	0.4754	0.1896	0.2268	Intimate	0.0504	0.0077	0.0627
Married	0.3452	0.1543	0.2397	Minority ethnic group	0.0502	-0.0122	0.0869
Paid work	0.3373	-0.1176	0.1866	Education level (2)	0.0490	-0.0068	0.0560
Voted last election	0.2367	0.0761	0.1565	Daily activities limitation	0.0469	-0.0044	0.0440
Borrow money	0.2338	0.0237	0.0489	Interested in politics	0.0466	-0.0029	0.0200
Left political orientation	0.2121	-0.0638	0.1415	Religious	0.0466	0.0009	0.0067
Age	0.1824	0.0019	0.0046	Income	0.0452	-0.0004	0.0095
Discriminated	0.1757	-0.1154	0.2906	Divorced	0.0448	0.0078	0.0729
Retired	0.1659	0.0529	0.1392	Supervising other employees	0.0434	-0.0019	0.0299
Widowed	0.1416	0.0830	0.2433	Number of household members	0.0395	-0.0002	0.0120
Country born	0.1329	0.0351	0.1069	Work from house	0.0393	0.0017	0.0306
Victim of burglary	0.1184	-0.0278	0.0908	Education job	0.0343	-0.0006	0.0998
Social activities	0.1144	0.0120	0.0404	Work hours	0.0314	0.0000	0.0007
Self-employed	0.0772	0.0222	0.0992	Member of a political party	0.0302	0.0008	0.0337
Attending religious services	0.0726	-0.0042	0.0195	Gender	0.0283	0.0000	0.0230
Separated	0.0723	-0.0405	0.1854	Unlimited job contract	0.0277	-0.0004	0.0314

Table 3.44: Bayesian Model Averaging Analysis: Netherlands

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Discriminated	1.0000	-0.5471	0.1107	Work hours	0.0516	-0.0002	0.0008
Income feeling	1.0000	-0.3984	0.0429	Minority ethnic group	0.0445	-0.0069	0.0412
General health	1.0000	-0.4937	0.0410	Separated	0.0376	-0.0152	0.1058
Religious	0.9987	0.0816	0.0169	Education level (2)	0.0356	0.0027	0.0189
Partner	0.9987	0.4103	0.0729	Age	0.0318	-0.0001	0.0006
Retired	0.997	0.3357	0.0779	Interested in politics	0.0302	0.0013	0.0107
Education level (1)	0.9785	0.3901	0.1184	Married	0.0261	-0.0014	0.0248
Social activities	0.9737	0.1280	0.0402	Pray	0.0248	0.0002	0.0026
Politics scale	0.9656	0.0649	0.0268	Income	0.0246	-0.0002	0.0027
Socializing	0.9381	0.0815	0.0314	Work from house	0.0243	0.0007	0.0110
Gender	0.3610	-0.0550	0.0814	Number of household members	0.0239	0.0005	0.0050
Voted last election	0.3437	0.0768	0.1183	Daily activities limitation	0.0232	-0.0006	0.0125
Victim of burglary	0.2706	-0.0495	0.0904	Living with children	0.0228	-0.0002	0.0109
Left political orientation	0.2026	0.0456	0.1020	Self-employed	0.0225	-0.0010	0.0171
Intimate	0.1455	0.0354	0.0971	Widowed	0.0222	-0.0008	0.0186
Borrow money	0.1438	0.0070	0.0195	Attending religious services	0.0206	-0.0002	0.0033
Education job	0.1144	0.0288	0.0917	Unemployed	0.0196	-0.0006	0.0251
Country born	0.0956	0.0183	0.0650	Member of a political party	0.0187	0.0005	0.0153
Divorced	0.0799	-0.0164	0.0656	Supervising other employees	0.0184	-0.0002	0.0087
Family-owned business	0.0597	0.0218	0.1065	Paid work	0.0182	0.0000	0.0121
Right political orientation	0.0529	-0.0041	0.0342	Unlimited job contract	0.0179	0.0004	0.0097

Table 3.45: Bayesian Model Averaging Analysis: Norway

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.4264	0.0433	Family-owned business	0.0609	0.0269	0.1257
Separated	1.0000	-1.1980	0.2258	Country born	0.0593	0.0105	0.0495
General health	1.0000	-0.4567	0.0362	Religious	0.0478	0.0010	0.0058
Social activities	1.0000	0.2710	0.0389	Income	0.0413	-0.0007	0.0044
Partner	0.9921	0.3141	0.0750	Number of household members	0.0401	0.0011	0.0074
Socializing	0.9756	0.0875	0.0264	Education job	0.0395	0.0057	0.0354
Retired	0.9739	0.3522	0.1003	Paid work	0.0378	-0.0045	0.0331
Intimate	0.8663	0.3941	0.1973	Member of a political party	0.0360	0.0040	0.0268
Politics scale	0.8388	0.0621	0.0379	Interested in politics	0.0283	0.0011	0.0093
Education level (2)	0.6837	0.1622	0.1273	Education level (1)	0.0277	0.0097	0.0835
Discriminated	0.3468	-0.1087	0.1651	Married	0.0261	0.0018	0.0208
Unemployed	0.3364	-0.1465	0.2282	Unlimited job contract	0.0258	0.0014	0.0134
Right political orientation	0.2772	-0.0662	0.1185	Living with children	0.0243	0.0005	0.0110
Left political orientation	0.1876	-0.0401	0.0908	Age	0.0217	0.0000	0.0005
Gender	0.1850	-0.0236	0.0553	Divorced	0.0207	0.0013	0.0180
Widowed	0.1752	0.0521	0.1266	Supervising other employees	0.0203	0.0006	0.0093
Daily activities limitation	0.1183	-0.0176	0.0545	Minority ethnic group	0.0187	-0.0016	0.0235
Pray	0.0978	0.0029	0.0103	Work hours	0.0179	0.0000	0.0003
Attending religious services	0.0755	-0.0035	0.0148	Voted last election	0.0177	-0.0008	0.0119
Work from house	0.0724	0.0077	0.0321	Self-employed	0.0162	-0.0004	0.0135
Borrow money	0.0649	0.0026	0.0115	Victim of burglary	0.0152	-0.0005	0.0090

Table 3.46: Bayesian Model Averaging Analysis: Poland

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-1.0031	0.0677	Married	0.0920	0.0129	0.0767
General health	1.0000	-0.6147	0.0493	Family-owned business	0.0771	0.0592	0.2359
Religious	1.0000	0.1689	0.0224	Attending religious services	0.0644	0.0048	0.0213
Social activities	0.9999	0.2239	0.0435	Member of a political party	0.0642	0.0314	0.1423
Intimate	0.9997	0.5840	0.1263	Work hours	0.0474	-0.0002	0.0010
Borrow money	0.9996	0.1980	0.0358	Age	0.0420	-0.0002	0.0014
Unemployed	0.9994	-0.7373	0.1421	Education level (2)	0.0388	0.0057	0.0354
Paid work	0.9993	-0.4237	0.0905	Work from house	0.0356	-0.0039	0.0278
Politics scale	0.9916	0.0895	0.0185	Socializing	0.0338	0.0011	0.0079
Income	0.9899	0.0767	0.0200	Self-employed	0.0335	0.0044	0.0311
Divorced	0.8431	-0.6306	0.3384	Living with children	0.0321	-0.0008	0.0213
Partner	0.6294	0.2012	0.1899	Voted last election	0.0319	0.0029	0.0222
Victim of burglary	0.5130	-0.1305	0.1424	Daily activities limitation	0.0305	-0.0026	0.0234
Discriminated	0.4873	-0.2423	0.2768	Left political orientation	0.0299	-0.0046	0.0393
Gender	0.3805	-0.0807	0.1145	Number of household members	0.0262	0.0004	0.0047
Separated	0.2992	-0.3517	0.5947	Minority ethnic group	0.0240	0.0045	0.0497
Pray	0.2762	0.0170	0.0305	Country born	0.0201	0.0036	0.0521
Interested in politics	0.1990	0.0224	0.0500	Supervising other employees	0.0180	-0.0007	0.0129
Education job	0.1672	0.0649	0.1619	Right political orientation	0.0178	0.0007	0.0216
Widowed	0.1604	0.0678	0.1764	Unlimited job contract	0.0164	-0.0008	0.0118
Education level (1)	0.1096	0.0750	0.2424	Retired	0.0150	-0.0004	0.0199

Table 3.47: Bayesian Model Averaging Analysis: Portugal

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.4651	0.0827	Work from house	0.0472	-0.0077	0.0483
Religious	1.0000	0.1961	0.0246	Socializing	0.0444	-0.0015	0.0099
Social activities	1.0000	0.4583	0.0632	Member of a political party	0.0413	-0.0103	0.0728
General health	0.9975	-0.3619	0.0778	Retired	0.0393	0.0056	0.0435
Borrow money	0.4736	0.0670	0.0795	Self-employed	0.0382	0.0049	0.0380
Gender	0.2745	0.0651	0.1183	Married	0.0367	0.0041	0.0322
Income	0.2616	0.0169	0.0319	Country born	0.0359	0.0073	0.0547
Daily activities limitation	0.2458	-0.0851	0.1673	Paid work	0.0341	-0.0039	0.0313
Education level (2)	0.2070	-0.0625	0.1378	Minority ethnic group	0.0320	-0.0089	0.0783
Discriminated	0.1558	-0.1073	0.2835	Education level (1)	0.0301	-0.0012	0.0277
Left political orientation	0.1302	-0.0254	0.0757	Widowed	0.0301	-0.0003	0.0283
Divorced	0.1126	-0.0458	0.1480	Unlimited job contract	0.0301	-0.0025	0.0228
Politics scale	0.1025	0.0069	0.0256	Living with children	0.0280	-0.0003	0.0175
Attending religious services	0.0984	0.0062	0.0216	Partner	0.0270	0.0010	0.0242
Voted last election	0.0843	0.0173	0.0677	Number of household members	0.0258	0.0001	0.0073
Interested in politics	0.0657	-0.0056	0.0255	Unemployed	0.0256	-0.0029	0.0377
Age	0.0635	-0.0004	0.0018	Family-owned business	0.0252	-0.0240	0.2738
Intimate	0.0621	0.0144	0.0703	Separated	0.0239	-0.0025	0.0645
Education job	0.0616	0.0401	0.1939	Victim of burglary	0.0210	-0.0002	0.0188
Right political orientation	0.0602	-0.0161	0.0860	Work hours	0.0202	0.0000	0.0003
Pray	0.0549	0.0019	0.0103	Supervising other employees	0.0190	0.0010	0.0191

Table 3.48: Bayesian Model Averaging Analysis: Russian Federation

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	0.9998	-0.6145	0.1237	Left political orientation	0.0623	-0.0157	0.0872
General health	0.9956	-0.6463	0.1519	Number of household members	0.0505	-0.0027	0.0209
Religious	0.9894	0.1764	0.0509	Interested in politics	0.0502	-0.0037	0.0283
Social activities	0.5707	0.1470	0.1465	Separated	0.0474	-0.0176	0.1467
Right political orientation	0.5684	0.2890	0.2906	Married	0.0471	-0.0048	0.0513
Discriminated	0.5022	-0.3614	0.4109	Paid work	0.0467	0.0086	0.0583
Intimate	0.4922	0.4124	0.4771	Divorced	0.0444	0.0094	0.0700
Victim of burglary	0.4477	-0.2556	0.323	Retired	0.0435	-0.0039	0.0497
Daily activities limitation	0.2776	-0.1314	0.2423	Minority ethnic group	0.0431	-0.0060	0.0570
Income	0.1972	0.0146	0.0341	Education level (1)	0.0429	0.0065	0.0824
Pray	0.1676	0.0135	0.0351	Partner	0.0429	-0.0061	0.0531
Supervising other employees	0.1389	0.0486	0.1433	Widowed	0.0427	-0.0011	0.0519
Borrow money	0.1335	0.0221	0.0672	Member of a political party	0.0418	-0.0090	0.1058
Socializing	0.1254	0.0120	0.0378	Work from house	0.0408	-0.0021	0.0521
Country born	0.1169	0.0722	0.2382	Education level (2)	0.0395	0.0031	0.0777
Attending religious services	0.0970	0.0112	0.0426	Age	0.0385	0.0000	0.0013
Politics scale	0.0897	0.0044	0.0256	Voted last election	0.0355	0.0001	0.0422
Work hours	0.0752	-0.0007	0.0035	Family-owned business	0.0349	0.0012	0.2461
Unlimited job contract	0.0710	0.0166	0.0834	Education job	0.0339	0.0034	0.1051
Self-employed	0.0664	-0.0399	0.2113	Living with children	0.0332	0.0006	0.0350
Gender	0.0623	-0.0117	0.0657	Unemployed	0.0332	-0.0037	0.0973

Table 3.49: Bayesian Model Averaging Analysis: Sweden

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.4799	0.0561	Pray	0.0568	-0.0011	0.0059
Retired	1.0000	0.6839	0.0949	Separated	0.0487	0.0278	0.1584
General health	1.0000	-0.5621	0.0551	Religious	0.0435	0.0006	0.0045
Socializing	0.989	0.1096	0.0296	Family-owned business	0.0435	0.0245	0.1515
Partner	0.9879	0.4047	0.0965	Voted last election	0.0396	0.0051	0.0340
Daily activities limitation	0.8554	-0.2418	0.1270	Divorced	0.0395	-0.0048	0.0330
Interested in politics	0.7708	0.1076	0.0704	Left political orientation	0.0391	0.0026	0.0331
Unemployed	0.7614	-0.4764	0.3186	Gender	0.0377	-0.0024	0.0180
Borrow money	0.7187	0.0668	0.0492	Education level (2)	0.0334	0.0020	0.0185
Social activities	0.6886	0.0842	0.0658	Widowed	0.0305	0.0040	0.0341
Attending religious services	0.3956	-0.0265	0.0365	Paid work	0.0291	0.0011	0.0214
Discriminated	0.2514	-0.0689	0.1332	Work from house	0.0280	0.0005	0.0230
Country born	0.2423	0.0578	0.1149	Unlimited job contract	0.0278	0.0004	0.0133
Politics scale	0.2041	0.0074	0.0169	Living with children	0.0271	0.0007	0.0141
Income	0.1454	-0.0064	0.0177	Member of a political party	0.0271	0.0014	0.0202
Minority ethnic group	0.1454	-0.0510	0.1418	Education job	0.0251	0.0009	0.0218
Education level (1)	0.1337	0.0280	0.0818	Victim of burglary	0.0234	-0.0003	0.0117
Right political orientation	0.1057	0.0129	0.0448	Number of household members	0.0224	0.0004	0.0058
Intimate	0.0966	0.0212	0.0764	Age	0.0222	0.0000	0.0004
Work hours	0.0707	-0.0003	0.0013	Supervising other employees	0.0221	0.0002	0.0109
Married	0.0617	0.0096	0.0475	Self-employed	0.0197	0.0007	0.0165

Table 3.50: Bayesian Model Averaging Analysis: Slovenia

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.6871	0.1064	Attending religious services	0.0746	-0.0053	0.0239
General health	0.9970	-0.4397	0.0966	Partner	0.0721	0.0102	0.0716
Income	0.9828	0.1512	0.0441	Minority ethnic group	0.0569	-0.0215	0.1237
Socializing	0.9783	0.1743	0.0516	Paid work	0.0566	0.0015	0.0727
Retired	0.8469	0.4864	0.2623	Voted last election	0.0559	0.0100	0.0580
Borrow money	0.7237	0.1383	0.1016	Victim of burglary	0.0540	-0.0105	0.0629
Discriminated	0.7203	-0.5875	0.4355	Family-owned business	0.0510	-0.0226	0.1432
Education level (2)	0.6539	-0.3317	0.2815	Gender	0.0484	-0.0052	0.0373
Right political orientation	0.4810	0.1800	0.2117	Work hours	0.0474	-0.0002	0.0015
Married	0.3164	0.1147	0.1949	Living with children	0.0458	0.0013	0.0415
Daily activities limitation	0.2920	-0.1112	0.1964	Member of a political party	0.0455	0.0084	0.0683
Intimate	0.2519	0.1224	0.2392	Divorced	0.0448	-0.0117	0.0947
Politics scale	0.1752	0.0111	0.0282	Interested in politics	0.0426	-0.0023	0.0204
Widowed	0.1543	0.0884	0.2460	Number of household members	0.0403	-0.0009	0.0121
Work from house	0.1494	0.0377	0.1045	Separated	0.0402	0.0125	0.1524
Unemployed	0.1210	-0.0571	0.1826	Left political orientation	0.0402	-0.0010	0.0427
Social activities	0.1177	0.0163	0.0529	Country born	0.0386	-0.0042	0.0491
Self-employed	0.1036	-0.0459	0.1615	Education job	0.0345	-0.0004	0.0421
Religious	0.0976	0.0045	0.0166	Unlimited job contract	0.0338	0.0017	0.0300
Education level (1)	0.0956	-0.0492	0.1857	Pray	0.0312	0.0001	0.0061
Age	0.0754	0.0006	0.0032	Supervising other employees	0.0311	0.0005	0.0256

Table 3.51: Bayesian Model Averaging Analysis: Slovak Republic

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.9884	0.1052	Work from house	0.0488	0.0079	0.0501
Religious	0.9992	0.1478	0.0323	Partner	0.0466	0.0072	0.0511
General health	0.9901	-0.3570	0.0991	Separated	0.0441	0.0224	0.1809
Unemployed	0.9653	-1.0163	0.3422	Pray	0.0440	0.0013	0.0102
Borrow money	0.9113	0.2223	0.0982	Married	0.0419	0.0060	0.0460
Victim of burglary	0.5283	-0.2813	0.3036	Social activities	0.0413	0.0012	0.0155
Family-owned business	0.5070	-1.2586	1.4132	Education job	0.0400	0.0039	0.0698
Socializing	0.2356	0.0216	0.0444	Intimate	0.0393	0.0049	0.0503
Unlimited job contract	0.2153	0.0909	0.2017	Daily activities limitation	0.0384	0.0027	0.0485
Self-employed	0.2078	0.1436	0.3254	Number of household members	0.0382	-0.0008	0.0098
Paid work	0.1187	-0.0342	0.1113	Interested in politics	0.0380	0.0012	0.0199
Age	0.0990	0.0008	0.0030	Supervising other employees	0.0369	-0.0035	0.0360
Politics scale	0.0792	0.0036	0.0155	Widowed	0.0364	-0.0049	0.0624
Right political orientation	0.0761	0.0166	0.0737	Divorced	0.0355	-0.0050	0.0677
Work hours	0.0742	-0.0004	0.0020	Left political orientation	0.0350	-0.0025	0.0350
Retired	0.0692	0.0171	0.0836	Gender	0.0337	-0.0002	0.0265
Discriminated	0.0671	-0.0297	0.1410	Education level (1)	0.0321	0.0080	0.1503
Education level (2)	0.0662	0.0163	0.0826	Minority ethnic group	0.0319	-0.0001	0.0525
Voted last election	0.0614	0.0112	0.0609	Attending religious services	0.0311	0.0001	0.0097
Country born	0.0527	0.0253	0.1513	Income	0.0288	-0.0003	0.0099
Living with children	0.0513	-0.0074	0.0479	Member of a political party	0.0274	0.0030	0.0828

Table 3.52: Bayesian Model Averaging Analysis: Turkey

Variable	PIP (i)	Post Mean (ii)	Post SD (iii)	Variable	PIP (iv)	Post Mean (v)	Post SD (vi)
Income feeling	1.0000	-0.9405	0.1244	Family-owned business	0.0525	0.0359	0.2381
General health	0.9989	-0.5496	0.1174	Country born	0.0522	-0.0422	0.2464
Right political orientation	0.8900	0.7328	0.3149	Paid work	0.0478	0.0060	0.0658
Borrow money	0.844	0.2272	0.1244	Voted last election	0.0419	-0.0075	0.0599
Discriminated	0.8377	-0.7989	0.4447	Minority ethnic group	0.0391	-0.0014	0.0801
Religious	0.4259	0.0435	0.0570	Separated	0.0383	-0.0225	0.2005
Unemployed	0.2394	-0.1326	0.2687	Partner	0.0369	-0.0090	0.1067
Education level (2)	0.1886	-0.0861	0.2045	Self-employed	0.0358	0.0001	0.0435
Socializing	0.185	0.0187	0.0452	Education level (1)	0.0347	-0.0022	0.0418
Victim of burglary	0.1677	-0.0815	0.2084	Married	0.0342	0.0054	0.1029
Gender	0.1480	-0.0604	0.1727	Left political orientation	0.0341	-0.0023	0.0554
Politics scale	0.1433	0.0155	0.0437	Pray	0.0339	0.0000	0.0122
Age	0.1177	0.0013	0.0042	Widowed	0.0338	0.0091	0.1013
Work from house	0.1013	0.0497	0.1787	Member of a political party	0.0334	-0.0051	0.0583
Attending religious services	0.0913	-0.0081	0.0318	Intimate	0.0326	0.0021	0.0344
Retired	0.0906	0.0348	0.1359	Unlimited job contract	0.0324	0.0024	0.0356
Income	0.0609	-0.0032	0.0165	Interested in politics	0.0320	0.0007	0.0150
Living with children	0.0556	-0.0094	0.0567	Education job	0.0311	0.0055	0.1028
Work hours	0.0554	-0.0003	0.0014	Daily activities limitation	0.0305	-0.0015	0.0447
Social activities	0.0532	0.0049	0.0300	Supervising other employees	0.0291	0.0028	0.0416
Number of household members	0.0526	-0.0032	0.0182	Divorced	0.0279	-0.0033	0.0951

Table 3.53: Bayesian Model Averaging Analysis: Ukraine

Variable	PIP	Post Mean	Post SD	Variable	PIP	Post Mean	Post SD
(i)	(ii)	(iii)	(iv)	(v)	(vi)		
Income feeling	1.0000	-0.9985	0.1673	Paid work	0.0738	0.0157	0.0983
General health	0.9837	-0.6337	0.1918	Education level (1)	0.0651	0.0302	0.1894
Victim of burglary	0.3935	-0.2863	0.4090	Attending religious services	0.0640	0.0054	0.0349
Family-owned business	0.3029	-0.6862	1.2041	Widowed	0.0619	0.0048	0.0858
Politics scale	0.2612	0.0277	0.0548	Unemployed	0.0601	-0.0131	0.1291
Divorced	0.1791	-0.1203	0.3081	Number of household members	0.0591	-0.0032	0.0248
Work from house	0.1518	-0.0621	0.1788	Separated	0.0583	-0.0209	0.2798
Social activities	0.147	0.0285	0.0840	Education level (2)	0.0569	-0.0099	0.1025
Discriminated	0.1182	-0.0835	0.2942	Member of a political party	0.0558	-0.0150	0.1324
Intimate	0.1122	0.0487	0.1745	Living with children	0.0556	0.0053	0.0610
Left political orientation	0.1042	-0.0304	0.1368	Pray	0.0533	0.0002	0.0143
Religious	0.1022	0.0071	0.0288	Gender	0.0527	-0.0036	0.0578
Country born	0.1019	-0.0469	0.1809	Supervising other employees	0.0521	-0.0053	0.0629
Socializing	0.1018	0.0097	0.0377	Borrow money	0.0517	0.0006	0.0307
Education job	0.0992	0.0791	0.3235	Interested in politics	0.0495	-0.0029	0.0359
Age	0.0954	0.0010	0.0041	Daily activities limitation	0.0482	0.0014	0.0653
Retired	0.0940	0.0302	0.1380	Income	0.0480	-0.0001	0.0127
Self-employed	0.0934	-0.0614	0.2596	Minority ethnic group	0.0475	-0.0009	0.0987
Unlimited job contract	0.0812	0.0294	0.1367	Work hours	0.0473	0.0001	0.0014
Voted last election	0.0798	0.0302	0.1465	Married	0.0461	-0.0015	0.0868
Right political orientation	0.0756	0.0107	0.1066	Partner	0.0458	-0.0023	0.0861

Chapter 4

Conditional Variance Decomposition

4.1 Abstract

A well established fact in the growth empirics literature is the increasing variation in output per capita across countries. This phenomena however does not adequately describe changes in the distribution of output since it does not account for changes in the covariates which undoubtedly influence per capita output levels. I propose a robust, nonparametric decomposition of the conditional variation of per capita output and find that OECD countries have experienced diminishing conditional variation while other regions have experienced increasing conditional variation. My decomposition suggests that most of these changes in the conditional variance of output is due to unobserved factors not accounted for by the traditional growth determinants. In addition to this I show as these factors played very different roles over time and across regions.

“An economically interesting formulation of conditional σ —convergence would be a useful contribution.” Durlauf et al. (2005, Page: 53)

4.2 Motivation

Cross country empirical growth studies commonly focus on β -convergence, in part to address such questions as “Do poor countries grow faster than richer ones?” or “How long will it take for a poor country to become rich?” Both of these questions are geared towards economies catching up with one another and highlight how relative income disparities are changing over time. However, it is well known that focusing on a coefficient in a conditional mean regression is limited (Quah 1993a) and cannot explain concepts such as intra-distributional churning, multi-modality, and expansion/contraction of the distribution over time. To more adequately study additional features of the cross-country distribution of output, growth empiricists have deployed a wide array of statistically rich modeling techniques to sharpen focus on how this distribution has changed. Within these studies a common ‘distributional moment’ that is of interest is the variance (see Pittau, Zelli & Johnson 2010), leading to speculation on σ -convergence.

In its most basic form, unconditional σ -convergence is assessed by looking at differences in the variation of the logarithm of cross-country output at two periods in time. As Quah (1996c) notes, while σ -convergence may be more illuminating regarding the behavior of the cross-country distribution of output than its β -convergence counterpart, it is still only a *feature* of the distribution and as such cannot capture entirely what is happening over time to the distribution. For instance, if one were to witness σ -convergence, intra-distribution churning and/or the appearance of multiple modes could occur, either of which would not be captured concomitantly with the observance of σ -convergence.

Given these shortfalls of studying different ‘moments’ of the distribution of output, one can still discern important information by studying the behavior of these moments over time. More directly, by focusing on the behavior of these moments in a conditional setting, empirical growth studies can glean information not available in the limited unconditional settings. Here, I use the methods similar to Maasoumi et al. (2007) to estimate the conditional density of cross-country per capita output. I also use the focus of Pittau et al. (2010) to analyze the variance of these conditional densities over time and for clusters of countries. The work of Maasoumi et al. (2007) focused primarily on the behavior of the conditional distribution/density of growth rates (actual and predicted) over time between OECD and non-OECD countries whereas the work of Pittau et al. (2010) decomposed the density of cross-country output into three groups and then subsequently analyzed the variance of these three groups over time. Here I blend these two studies together and offer a decomposition of the variance of the conditional density¹. I drift from Maasoumi et al. (2007) by explicitly focusing on the variance of the distribution while I extended the work of Pittau et al. (2010) by examining conditional variances as opposed to unconditional ones. In addition to this I try to decompose the relative weights in these variance changes of covariates and other effects, in a logic that is close to Beaudry, Collard & Green (2005). To obtain a sense for the importance of conditioning, consider the multimodality finding of Quah (1993a). While subsequently illuminating regarding the relative polarization of the distribution over

¹My clusters are defined according to a geographical criterion, but are very similar to clusters determined according to the location inside the distribution of output, given the income differences among continents.

time, given the unconditional framework of the analysis, only speculative evidence could be provided underlying the root for the emerging bimodal shape and increasing variance. The work of Caselli & Feyrer (2007) and Henderson, Parmeter & Russell (2008) both examined the behavior of the corresponding Solow growth determinants, along with cross-country output to see if similar patterns emerged that may provide further evidence to Quah's initial discovery of a bimodal distribution.

Similarly, this paper focuses on the conditional variation of the density of cross-country output over time to discern if any of the Solow growth determinants behavior corresponds to the increase in variation we have witnessed over the past 5 decades. A broad description of this in Table 4.1 provides summary variation measures for the 98 countries that have the appropriate data in Penn World Table 6.3 and Barro & Lee (2010) from 1960 to 2000. We immediately notice that all three of the common measures of per capita output (RGDPL, RGDPCH and RGDPWOK) display increasing variance over each of the decades. While this yields conclusions regarding the lack of σ -convergence, the underlying causes for this increase remain illusive in an unconditional setting². If we expand our focus to variation changes in the so called 'Solow' growth determinants, investment rates (INV), population growth (POP) and years of schooling (EDUC), we see that the variation in investment rates seems to be declining over time, in line with the research of Caselli & Feyrer (2007) while cross-country levels of education do not display absolute σ -convergence. The apparent σ -divergence of years of education is in accord with the development accounting findings of Stamatakis & Petrakis (2005) and Henderson et al. (2008).

Table 4.1: Variance of Logarithm of Output and Growth Accounting Variables Over Time.

	RGDPL	RGDPCH	RGDPWOK	INV	POP	EDUC
	(i)	(ii)	(iii)	(iv)	(v)	(vi)
1960	0.893	0.887	0.954	0.0149	0.0183	6.21
1970	0.985	0.980	1.014	0.0152	0.0176	7.41
1980	1.093	1.089	1.047	0.0131	0.0183	8.50
1990	1.248	1.244	1.119	0.0116	0.0238	8.42
2000	1.531	1.524	1.351	0.0104	0.0178	8.44

Notes: This table provides summary variation measures for the 98 countries that have the appropriate data in Penn World Table 6.3 and Barro & Lee (2010) from 1960 to 2000. Columns (i), (ii), and (iii) show the variance of logarithm of real per capita chain output series, real per capita output Laspeyres, real per worker output, respectively. Columns (iv), (v), and (vi) show the variance of growth accounting variables, physical capital, population growth, and human capital, respectively.

An interesting aspect of considering variation of conditional moments over time is how they correspond to the common notions of conditional convergence prevailing in the growth em-

²I caution that the appearance of σ -divergence can be attributed to the measure of dispersion used and not the actual phenomena of output diverging over time (Carl-Johan Dalgaard & Vastrup 2001). If we use the coefficient of variation instead of the standard deviation we find substantial stability for the period 1960-2000.

pirics literature currently. In the β -convergence literature the conditioning was designed to account for differences in the steady state levels of output across countries. However, when one migrates from a first moment setting to the investigation of higher moments, such as the conditional variance, it is not clear what conditional convergence of a ‘moment’ implies. While researchers have offered regression based tests of both unconditional and conditional σ -convergence (Cannon & Duck 2000, Egger & Pfaffermayr 2009, Huber & Pfaffermayr 2010), as noted by Bliss (2000), the fact that the variance of the left hand side variable is changing over time renders asymptotic analysis of the test difficult to understand. Broadly, this criticisms passes to any type of regression (conditional mean) setting which analyzes convergence of higher order moments (conditional or not). More importantly, Durlauf et al. (2005) have raised several issues with constructing suitable notions for conditional σ -convergence. Rather, to crystallize my focus I simply point out that instead of focusing on convergence or divergence, I instead seek to understand how conditional moments change over time, irrespective of whether they may be termed to conditionally diverge or converge.

It is useful to discuss what a conditional measure of variance should look like. As noted by Quah (1993*b*, 1993*a*), regressions that are done over time can fail to capture the underlying distribution dynamics, resulting in researchers misinterpreting their results, or, more starkly, for the results to be meaningless entirely. In fact, the notion of unconditional σ -convergence can be seen as the variance from the unconditional density of the logarithm of output at two points in time. Similarly, one may then suggest that the appropriate concept of conditional σ -convergence be defined by looking at the variance of the conditional density at two points in time. The key distinction here though, however, is that now that conditioning variables are present one needs to define exactly how the two variances are compared.

Taking this notion further, in a β -convergence study (either absolute or conditional) a regression of growth rates over time is run on some initial value of output or income and ‘conditioning’ variables. A negatively signed coefficient estimate on the initial income variable is taken as evidence for conditional β -convergence. The definition of absolute σ -convergence however, is a relationship between the variance of the logarithm of output at *two* points in time. Thus, one should not be looking to run a regression where a negative coefficient on some level of income at a point in time signifies that the conditional variance has decreased.³ While it is appealing to focus on aspect of ‘convergence’ in growth empirics, it also draws interest away from the appropriate settings to study the behavior of conditional moments of the distribution of cross-country output. Without formal links to the variables which influence cross-country output, it is hard to attach any meaning to why we see increased dispersion of incomes over time, let alone lay claim that these increases are interpretable from the standpoint of a specific variable. Again, I stress that looking at conditional densities over time is a more parsimonious approach to studying the behavior of a conditional moment (such as the variance) than unconditional, and subsequently, regression based conditional approaches.

³In addition to this, if the shape of the error distribution at two points of time is different, for example normally distributed at time t but a bimodal density at time T , regression-based tests for conditional σ -convergence will not be informative and most likely will use incorrect asymptotic distributions as the basis for constructing p -values.

If I define unconditional σ -convergence as

$$\sigma_{Y,t}^2 > \sigma_{Y,T}^2, \quad \text{for } T > t, \quad (4.1)$$

the key issue is that regardless whether or not we witness a diminution or an expanse of the variance over time, we cannot detect substantive links without appropriate conditioning. Thus, an appropriate measure of assessing the behavior of the conditional variation over time is

$$\sigma_{Y_i|X_i,t}^2 > \sigma_{Y_i|X_i,T}^2, \quad \text{for } T > t, \quad (4.2)$$

where the conditioning takes place at the same level of the covariates across time. Thus, what I am looking at *within* the conditional density of the logarithm of output is, for a given level of the covariates over time, has the dispersion of the log of output diminished. Thus, it is possible that the conditional variance is diminishing over time in certain covariate regions but not at others.⁴ of countries but not between groups. This is inline with recent, robust conditional β -convergence studies that incorporate interactions and nonlinearities that involve the initial value of income (Durlauf & Johnson 1995, Liu & Stengos 1999, Durlauf, Kourtellos & Minkin 2001, Kourtellos 2003). Thus, I can ask questions like “If African nations had levels of human capital and population growth as in OECD countries, would we witness a diminution of income dispersion over time?” or more generally “Without the observed changes in human or physical capital stocks would I observed less dispersion”. These questions are undoubtedly more interesting than their unconditional counterparts and also provide valuable policy insights.

The remainder of the paper is organized as follows. The second section outline my construction of the conditional density and how I will estimate the conditional variance. Then, I will investigate whether or not conditional σ -convergence is apparent in my data using a traditional Mankiw, Romer & Weil (1992) type dataset. Final section offers policy insights and avenues for further research.

4.3 Empirical Methods

While many growth studies begin with a regression framework, here I focus on the insights of Quah (1993*b*, 1993*a*) and use a distributional approach. Previous work that has studied various features of the growth process through a density (distribution) framework are Quah (1996*a*, 1996*b*, 1996*c*), Bianchi (1997), Jones (1997), Pritchett (1997), Paapaa & Van Dijk (1998), Desdoigts (1999), Johnson (2000), Gisbert (2003), Anderson (2004), Azariadis & Stachurski (2003), Canova (2004), Fiaschi & Lavezzi (2003), Johnson (2005), Pittau (2005), Pittau & Zelli (2006), Maasoumi et al. (2007), Henderson et al. (2008), and Pittau et al. (2010). Some of the papers have focused on unconditional density features while others have focused on conditional densities. Regardless, this list shows the growing body of literature that attempts to learn about the growth process via distributional methods. I introduce a generalized product kernel approach to constructing the conditional density over time that

⁴Here group is defined by similarities in the conditioning variables.

has not previously been used in the growth empirics literature.⁵

4.3.1 Estimating a Conditional Density

Let $f(\cdot)$ denote the joint density of (X, Y) and $g(\cdot)$ denote the marginal density of X . Here Y is the logarithm of output and X are the set of conditioning variables that will form the basis of the conditioning set to study conditional σ -convergence. Define the generalized product kernel as

$$K_h(x, X_i) = \prod_{s=1}^q h_s^{-1} k^c \left(\frac{x_s^c - X_{si}^c}{h_s} \right) \prod_{s=1}^r k^u(x_s^u, X_{si}^u, \lambda_s^u) \prod_{s=1}^p k^o(x_s^o, X_{si}^o, \lambda_s^o). \quad (4.3)$$

Here, X is partitioned into three components, continuous (X^c), ordered discrete (X^o), and unordered discrete (X^u). The distinction between data types is important in nonparametric applications as it lessens the curse of dimensionality. I have used q , r and p to denote the number of continuous, unordered, and ordered conditioning variables, respectively. $h_s^{-1} k^c(\cdot)$ is the standard normal kernel function with window width $h_s^c = h_s(N)$ associated with the s^{th} component of x^c that is commonly used in unconditional density studies (Bianchi 1997, Jones 1997, Pritchett 1997). k^u is a variation of Aitchison & Aitken's (1976) kernel function which equals one if $x_{si}^u = x_{sj}^u$ and λ_s^u otherwise, and k^o is the Wang & Van Ryzin's (1981) kernel function which equals one if $x_{si}^o = x_{sj}^o$ and $(\lambda_s^o)^{|x_{si}^o - x_{sj}^o|}$ otherwise. h^c , λ^u and λ^o are the bandwidths associated with each kernel. See Li & Racine (2003) and Li & Racine (2007) for further details.

I denote my estimators of the joint and marginal densities, $\hat{f}(x, y)$ and $\hat{g}(x)$ as

$$\hat{f}(x, y) = n^{-1} \sum_{i=1}^n K_h(x, X_i) k_{h_y}(y, Y_i), \quad (4.4)$$

$$\hat{g}(x) = n^{-1} \sum_{i=1}^n K_h(x, X_i), \quad (4.5)$$

where h_y is the smoothing parameter associated with Y and $k_{h_y}(\cdot)$ is the simple univariate, continuous kernel which smooths my outcome variable. Noting that the conditional density of Y is defined as $m(y|x) = f(x, y)/g(x)$, I estimate the conditional density as

$$\hat{m}(y|x) = \hat{f}(x, y)/\hat{g}(x). \quad (4.6)$$

Once I have obtained an estimate of the conditional density I can construct the conditional variance of the density for any level of the covariates ($x = \bar{x}$, say). Thus, I can estimate the conditional density at two points in time and hold the covariates fixed at their initial time period levels to determine if the conditional variance of y has decreased. All that remains is to discuss how the bandwidths, which are an integral part of the conditional density

⁵An exception is Maasoumi et al. (2007), but they use a generalized product kernel in a regression setting.

estimator, are arrived at and how I choose to estimate the variance at any particular level of covariates.

4.3.2 Bandwidth Selection

Bandwidth selection is commonly perceived as the most important aspect of any kernel based nonparametric modelling. While optimality results exist concerning the size of the bandwidths, I advocate a data-driven approach which has recently been hailed for its desirable asymptotic properties (Hall, Racine & Li 2004) and ability to remove the inclusion of irrelevant variables. The approach is termed Least Squares Cross-Validation (LSCV). The bandwidths are selected by minimizing integrated squared error:

$$ISE = \int \{\hat{m}(y|x) - m(y|x)\}^2 g(x) dx dy. \quad (4.7)$$

ISE can be written as the sum of three components ($ISE_1 + ISE_2 + ISE_3$), only two (ISE_1 and ISE_2) of which depend on the unknown bandwidths. Formally, I have

$$ISE_1 = n^{-1} \sum_{i=1}^n \hat{G}_{-i}(X_i) / \hat{g}_{-i}(X_i)^2, \quad (4.8)$$

$$ISE_2 = n^{-1} \sum_{i=1}^n \hat{m}_{-i}(X_i, Y_i) / \hat{g}_{-i}(X_i)^2, \quad (4.9)$$

where

$$\hat{G}(X_i) = n^{-2} \sum_{i_1=1}^n \sum_{i_2=1}^n K_\gamma(X_i, X_{i_1}) K_h(X_i, X_{i_2}) \int k_{h_y}(y, Y_{i_1}) k_{h_y}(y, Y_{i_2}) dy, \quad (4.10)$$

and a subscript $-i$ denotes a leave one out estimator, for example,

$$\hat{g}_{-i}(X_i) = (n-1)^{-1} \sum_{j=1, j \neq i}^n K_h(X_i, X_j). \quad (4.11)$$

The LSCV objective function is given as

$$CV(h_y, \gamma) = \widehat{ISE}_1(h_y, \gamma) - 2\widehat{ISE}_2(h_y, \gamma), \quad (4.12)$$

where $\gamma = \{h_1, \dots, h_q, \lambda_1^o, \dots, \lambda_p^o, \lambda_1^u, \dots, \lambda_r^u\}$ is the vector of bandwidths associated with the covariates. Once the bandwidths have been determined the conditional density can be estimated at any given level of the covariates.

4.3.3 Constructing the Variance

Constructing moments from a conditional density estimator is common. For example, when one estimates a regression model, this is the mean from the conditional density of the independent and dependent variables, i.e.,

$$\hat{E}[Y|X = x] = \int y \hat{m}(y|X = x) dy = \sum_{i=1}^n w_i(x) Y_i, \quad (4.13)$$

where $w_i(x) = K_h(x, X_i) / \sum_{i=1}^n K_h(x, X_i)$, which is the typical local constant kernel regression estimator of Nadaraya (1964) and Watson (1964). In my setting I am interested in the conditional variance. A little bit of algebra reveals that the estimate of the conditional variance (when Y is continuous), $\widehat{Var}[Y|X = x]$, is⁶

$$\hat{v}(x) = \int \left[y - \hat{E}[Y|X = x] \right]^2 \hat{m}(y|X = x) dy = h_y \sigma_K^2 + \sum_{i=1}^n w_i(x) \left[Y_i - \hat{E}[Y|X = x] \right]^2. \quad (4.14)$$

Here, σ_K^2 is the variance of the kernel used to smooth Y . In my empirical example I use a second order Gaussian kernel, $K(x) = h^{-1}(x/h)$, which has unit variance. I eschew integration by noting that my conditional variance can be calculated in three steps. First, conduct least squares cross validation to determine the optimal bandwidths for the conditional density. Second, estimate the conditional mean of Y using the bandwidths corresponding to the X s. Note that h_y is not used in this second stage. Third, using the estimates of the conditional mean obtained in the second stage, regress the squared residuals on X using the same set of bandwidths as those used for the conditional mean.

To investigate conditional σ -convergence I can look at the difference between my conditional variance estimates. That is,

$$\begin{aligned} \hat{v}_t(x_t) - \hat{v}_T(x_T) &= \underbrace{\hat{v}_t(x_t) - \hat{v}_T(x_t)}_{\text{Time Effect}} + \underbrace{\hat{v}_T(x_t) - \hat{v}_T(x_T)}_{\text{Covariate Effect}} \\ &= \underbrace{\hat{v}_t(x_t) - \hat{v}_t(x_T)}_{\text{Covariate Effect}} + \underbrace{\hat{v}_t(x_T) - \hat{v}_T(x_T)}_{\text{Time Effect}}. \end{aligned} \quad (4.15)$$

My concern is with the covariate effect. It tells me the contribution to the overall change in the conditional variance based on how the covariates have changed over time, holding the shape of the distribution fixed. Thus, in order to say that the variance went down, the change in the covariates must have moved the country in such a way that the conditional variance decreased, which suggests conditional variance diminishes relative to the covariates. As an example, if a country's covariates changed over time so that there was massive population growth, no human capital and the investment rate was zero, this should cause the variation in potential output to all but disappear because these levels of the covariates

⁶See Hyndman, Bashtannyk & Grunwald (1996) for more on this derivation in the simple, univariate kernel setting.

in unison would not promote growth, no matter the shocks. The time effect is how the conditional variance function itself has changed over time. It is linked to all other things but growth determinants stocks. Two obvious candidates are TFP and coefficients (returns) of growth factors (this way is linked to a traditional Oaxaca type decomposition). In addition to this, if time dependent heteroscedasticity were present this term could trigger either the appearance of σ -convergence or divergence if it were large enough in magnitude.

This point is further discussed in Durlauf et al. (2005). Thus, by focusing on specific levels of covariates I can determine if their shift has been promoting a collapse of the variance of the conditional density over time or not. If the covariate effect is zero then I can conclude that any increase/decrease is due to some underlying time varying factors. For instance it could mean that with comparable productive stocks an asymmetric change in the production function played a big role in expanding income variance over time. However, if the covariate effect is not zero then this tells me that the impact on the dispersion of output depends on where a country lies within the joint distribution. Thus, policy makers can use this information to dictate how they choose to allocate resources to promote different types of factor accumulation for different countries. Due to the fully nonparametric setup of my problem, the choice of covariate levels in both time periods will dictate the appearance of conditional σ -convergence. However, it is not obvious which set(s) of points to look for this pattern of conditional convergence.

Another way to look at (15) is like a distributional generalization of the decomposition approach of Beaudry et al. (2005) where they build counterfactual income distributions by growth rate accumulation according to Barro type regressions. That approach implies an homogeneity assumption of production function across countries, given by the fact that they apply the same coefficients to each part of income distribution (so that a shift in mean is representative for all the distributional shifts). In addition to this they have to accumulate growth rates in order to avoid endogeneity problems that could be caused by having income level on the left hand side. In my case the kernel approach showed in (13) preserves me by this restrictive assumption and allows to use directly income levels.

4.4 Empirics

4.4.1 Data

The data used for this study come from the Penn World Table, Mark 6.3 (Heston, Summers & Aten 2009). I used RGDPCH to measure output per person, and I construct population growth rates and physical capital stocks from the corresponding variables in the PWT. My measure of human capital stocks are the average years of secondary schooling for the population from Barro & Lee (2010).

Before determining the change of the conditional variance over time I look at the basic properties of the unconditional variables over time, in order to understand if they show asymmetric convergence processes within geographical groups and over time as well. The changes are relative to each variable so that they may be compared. Following equation

(4) in Mankiw et al. (1992) I look at the logarithm of each variable and given that in the following decomposition I will implement also a sort of income variance accounting exercise. Table 4.2 shows the results for a sample of 98 countries for the period 1960-2000. I may see as the results are very heterogenous by time, groups and variable⁷. There is evidence of global σ -divergence for GDP while only OECD shows sustained convergence. Interestingly, human capital and physical capital investment rates have opposite trends with the former having divergence in 70s driven by Asian and African countries (that have high divergence inside groups also in the following decades), while the latter being converging after 70s with the exception of the Asian group in the 80s and of the LAC group in the 90s⁸. Finally, growth population variance has an increasing pattern in the 80s and then decreasing in the 90s.

4.4.2 Results

Using the methods described in Section 4.3, I estimate the conditional distribution of output based on 10 years averages of population growth, human capital stocks and investment rates. I do this for the years 1960 and 2000 and then construct my measure of conditional σ -convergence accordingly. Since I am subtracting year 2000 numbers from year 1970 numbers, when $\hat{v}_{1970}(x_{1970}) - \hat{v}_{2000}(x_{2000})$ is positive, this implies conditional σ -convergence. A negative value thus signifies divergence.

For ease of discussion I first discuss conditional σ -convergence for each country, grouping them regionally. I then discuss conditional σ -convergence on the mean and quartile values of the covariates across time. I split the period into two waves: 1960-1980 and 1980-2000, in order to see what happens before and after the oil shocks and the productivity slowdown. I begin with country specific measures. The OECD country results are presented in Table 4.3. I notice three distinct features. First, the majority of these countries have experience conditional convergence of the variance. Second, aside from Mexico and Turkey, every country witnessed a positive time effect. This time effect captures time dependent heteroscedasticity is the reason that any OECD country has experienced conditional σ -convergence. Third, the covariate effect is negative (but smaller) for every country, suggesting that the change in investment rates, human capital accumulation and population growth contributed to increase of variation in output. Fourth, the great bulk of conditional convergence is in the first period of time, while after 1980 the situation is stable and there is a reversion of time effects.

Given that recent studies have highlighted the difference between African countries and the rest of the world, I isolate the Sub-Saharan countries within my dataset in Table 4.4. The results are almost identical to those for the OECD countries except in one facet, almost all African nations experience conditional σ -divergence. In this case the covariate effects are all positive except for four countries, while for the most part of the countries, the time

⁷A decreasing pattern inside each group could be still consistent with an increasing variance if the distribution becomes more multimodal, so if there is a divergent(across)-convergence(within) phenomenon.

⁸In a multivariate clustering framework Battisti & Parmeter (2010) found asymmetries among relationships of GDP with physical and human capital stocks among richer and poorer groups.

Table 4.2: Relative Changes in Variances of Logarithm of Output and Growth Accounting Variables Over Time by Geographical Groups 1960-2000.

	1960-1970	1970-1980	1980-1990	1990-2000
RGDPWOK	(i)	(ii)	(iii)	(iv)
All	0.105	0.031	0.069	0.187
OECD	-0.378	-0.095	-0.122	0.077
South American	0.389	-0.254	0.002	0.341
Asian	0.218	0.362	0.209	0.117
African	0.124	0.113	0.163	0.162
Mena	-0.114	-0.413	0.107	0.252
POPULATION GROWTH				
All	-0.030	0.038	0.298	-0.251
OECD	-0.349	-0.221	-0.092	-0.410
South American	0.315	-0.224	0.172	-0.215
Asian	-0.848	0.194	1.077	0.435
African	-0.009	0.189	0.399	-0.317
Mena	0.395	-0.305	0.050	0.086
YEARS OF EDUCATION				
All	0.192	0.146	-0.013	0.001
OECD	0.075	0.038	-0.254	-0.254
South American	0.451	-0.104	0.161	-0.053
Asian	0.391	0.387	0.302	0.198
African	0.301	0.221	0.790	0.378
Mena	0.215	0.260	-0.089	-0.164
INVESTMENT RATES				
All	0.048	-0.139	-0.119	-0.102
OECD	0.100	-0.163	-0.335	-0.392
South American	-0.178	-0.044	-0.529	0.415
Asian	0.392	-0.271	0.238	-0.337
African	-0.294	-0.208	-0.057	-0.463
Mena	0.296	-0.315	-0.247	-0.456

Notes: This table looks at the basic properties of the unconditional growth accounting variables over time, in order to understand if they show asymmetric convergence processes within geographical groups and over time as well. These variables are real per worker output (RGDPWOK), population growth, years of education (human capital), and investment rates (physical capital). The geographical regions are listed in the most left column and columns (i)-(iv) show the time intervals.

effects are negative. On average the measure of divergence is quite large. These results are suggestive that even if Africa had the factors of today back in 1970, the persistence of shocks that appear to have increased over time, has created a lack of convergence of the variance over time. In fact if I look the two subperiods, I have that almost all the divergence is in the second period with a highly negative covariate effect.

Tables 4.5 through 4.7 present analogous results for the Latin American countries (LAC), the Asian countries and the group of countries that comprise the Middle East and North Africa

Table 4.3: Overall, Time and Covariate Effects for Conditional Variance for OECD Countries

	1960-2000			1960-1980			1980-2000		
Country	Overall	Covariate	Time	Overall	Covariate	Time	Overall	Covariate	Time
Australia	-0.005	-0.071	0.066	0.011	-0.046	0.058	-0.016	0.009	-0.026
Austria	-0.003	-0.155	0.152	0.010	-0.156	0.166	-0.013	-0.002	-0.011
Belgium	0.051	-0.107	0.158	0.056	-0.061	0.117	-0.005	0.001	-0.006
Canada	-0.033	-0.113	0.080	-0.026	-0.115	0.089	-0.007	0.014	-0.021
Denmark	0.077	-0.066	0.143	0.087	-0.024	0.111	-0.010	-0.005	-0.004
Finland	-0.001	-0.119	0.118	-0.003	-0.139	0.136	0.002	0.005	-0.003
France	-0.001	-0.150	0.149	-0.003	-0.095	0.092	0.002	0.012	-0.010
Greece	0.137	-0.004	0.141	0.118	0.049	0.069	0.019	0.023	-0.004
Iceland	-0.031	-0.161	0.130	-0.063	-0.140	0.077	0.031	0.024	0.007
Ireland	0.000	-0.142	0.142	-0.074	-0.119	0.046	0.074	0.075	-0.001
Italy	0.016	-0.142	0.158	0.008	-0.087	0.095	0.008	0.013	-0.006
Japan	0.126	-0.043	0.168	0.072	-0.027	0.099	0.054	0.061	-0.007
Luxembourg	0.052	-0.009	0.061	0.086	-0.040	0.127	-0.034	-0.027	-0.007
Mexico	-0.024	-0.003	-0.021	-0.058	-0.054	-0.004	0.034	0.062	-0.028
Netherlands	-0.004	-0.144	0.140	-0.003	-0.136	0.133	-0.001	0.012	-0.012
New Zealand	0.002	-0.096	0.098	0.022	-0.073	0.094	-0.019	-0.008	-0.011
Norway	-0.008	-0.154	0.146	-0.012	-0.116	0.104	0.004	0.014	-0.011
Portugal	0.037	-0.061	0.099	0.03	0.027	0.003	0.008	0.007	0.001
Spain	0.051	-0.109	0.160	0.029	-0.015	0.043	0.022	0.027	-0.005
Sweden	0.089	-0.058	0.147	0.096	-0.047	0.143	-0.007	0.002	-0.009
Switzerland	0.127	-0.021	0.148	0.144	-0.030	0.174	-0.018	-0.007	-0.011
Turkey	-0.141	-0.114	-0.026	-0.017	-0.029	0.011	-0.123	-0.089	-0.035
United Kingdom	0.009	-0.103	0.112	0.027	-0.040	0.067	-0.018	-0.010	-0.008
United States	0.030	-0.027	0.057	0.063	-0.031	0.094	-0.033	-0.004	-0.028
Average	0.023	-0.091	0.114	0.025	-0.064	0.089	-0.002	0.009	-0.011

Notes: This table provides the estimation of the conditional distribution of output based on 10 years averages of population growth, human capital stocks and investment rates for the OECD countries. This is done for the years 1960 and 2000 and then the measure of conditional σ —convergence is constructed accordingly. The results are divided to overall effect, covariate effect, and time effect.

(MENA). Again, the majority of these countries have witnessed conditional σ -divergence, with mostly negative time effects. Of the four groups of countries, as expected Africa has the highest average of conditional σ —divergence while Asia has the lowest. In the whole period, Asian, LAC and MENA countries have similar low average rates of divergence with respect to Africa. Looking to the time pattern highlights interesting findings: the big Asian divergence of 1960-1980 was partially balanced by the lower convergence of 1980-2000, while for LAC there had been a positive covariate effect in 1980-2000 that had been offset by the

Table 4.4: Overall, Time and Covariate Effects for Conditional Variance for African Countries

	1960-2000			1960-1980			1980-2000		
Country	Overall	Covariate	Time	Overall	Covariate	Time	Overall	Covariate	Time
Benin	-0.300	0.051	-0.351	0.031	0.010	0.022	-0.332	-0.340	0.008
Botswana	-0.237	0.132	-0.369	0.122	-0.003	0.124	-0.358	-0.308	-0.050
Burundi	0.070	0.029	0.041	0.061	-0.039	0.100	0.009	0.017	-0.008
Cameroon	-0.257	0.118	-0.375	0.002	0.113	-0.110	-0.260	-0.296	0.036
Cen. African	-0.252	0.088	-0.340	0.095	0.054	0.041	-0.347	-0.305	-0.042
Congo Dem.	-0.301	0.102	-0.404	0.101	0.062	0.039	-0.402	-0.313	-0.090
Congo Rep.	-0.146	0.035	-0.181	-0.250	0.038	-0.289	0.104	-0.297	0.401
Cote Divoire	0.079	0.056	0.023	0.010	-0.036	0.047	0.068	0.053	0.015
Gabon	-0.487	0.107	-0.594	-0.204	0.109	-0.312	-0.283	-0.248	-0.035
Gambia	-0.245	0.039	-0.284	-0.002	0.015	-0.018	-0.242	-0.298	0.055
Ghana	-0.670	-0.019	-0.650	-0.443	-0.018	-0.425	-0.227	-0.152	-0.075
Kenya	-0.374	0.035	-0.410	-0.023	-0.045	0.022	-0.351	-0.255	-0.096
Lesotho	0.114	0.102	0.012	0.139	0.066	0.073	-0.025	-0.091	0.066
Malawi	0.038	0.094	-0.057	0.102	-0.005	0.108	-0.065	-0.062	-0.002
Mali	0.067	0.007	0.060	0.057	-0.026	0.083	0.009	0.000	0.009
Mauritania	-0.232	0.066	-0.298	0.033	0.020	0.013	-0.265	-0.152	-0.113
Mauritius	-0.388	0.006	-0.393	-0.181	0.090	-0.271	-0.207	0.094	-0.300
Morocco	-0.027	-0.047	0.021	-0.043	-0.043	0.000	0.016	-0.034	0.050
Mozambique	0.057	0.023	0.035	0.064	-0.063	0.126	-0.006	-0.054	0.048
Namibia	-0.430	0.004	-0.433	-0.099	-0.056	-0.043	-0.331	-0.286	-0.045
Niger	0.082	0.060	0.022	0.103	-0.004	0.108	-0.021	-0.050	0.029
Rwanda	0.047	-0.011	0.058	0.025	-0.003	0.028	0.022	0.016	0.007
Senegal	-0.214	0.128	-0.342	0.129	0.115	0.014	-0.342	-0.291	-0.052
South Africa	-0.444	-0.012	-0.432	-0.239	-0.008	-0.231	-0.205	-0.151	-0.054
Tanzania	0.061	0.067	-0.006	0.063	0.070	-0.006	-0.002	-0.006	0.004
Togo	-0.317	0.094	-0.412	0.024	0.015	0.009	-0.341	-0.236	-0.105
Uganda	0.003	0.073	-0.070	0.083	-0.008	0.091	-0.080	-0.176	0.096
Zambia	-0.272	0.133	-0.405	0.112	0.051	0.061	-0.385	-0.247	-0.137
Zimbabwe	-0.343	0.028	-0.371	-0.028	0.100	-0.129	-0.314	-0.049	-0.265
Average	-0.183	0.055	-0.238	-0.018	-0.025	0.008	-0.178	-0.156	-0.022

Notes: This table provides the estimation of the conditional distribution of output based on 10 years averages of population growth, human capital stocks and investment rates for the African countries. This is done for the years 1960 and 2000 and then the measure of conditional σ -convergence is constructed accordingly. The results are divided to overall effect, covariate effect, and time effect.

time.

Table 4.5: Overall, Time and Covariate Effects for Conditional Variance for Latin American Countries

	1960-2000			1960-1980			1980-2000		
Country	Overall	Covariate	Time	Overall	Covariate	Time	Overall	Covariate	Time
Argentina	0.031	0.051	-0.020	0.060	0.033	0.026	-0.028	0.031	-0.059
Barbados	-0.078	-0.133	0.055	-0.058	-0.034	-0.024	-0.021	0.034	-0.054
Bolivia	-0.074	0.021	-0.095	0.030	0.011	0.019	-0.103	0.006	-0.109
Brazil	-0.109	-0.031	-0.078	-0.042	-0.065	0.023	-0.067	0.048	-0.115
Chile	-0.012	-0.009	-0.003	0.015	-0.001	0.016	-0.027	0.029	-0.055
Colombia	-0.112	-0.011	-0.101	-0.017	-0.067	0.050	-0.095	0.008	-0.103
Costa Rica	-0.045	0.004	-0.049	-0.061	-0.061	0.000	0.016	0.070	-0.054
Dominican Republic	-0.070	-0.044	-0.027	-0.097	-0.110	0.013	0.026	0.046	-0.020
Ecuador	-0.049	0.016	-0.065	-0.018	0.020	-0.038	-0.031	0.061	-0.092
El Salvador	-0.161	-0.092	-0.069	-0.034	-0.044	0.009	-0.127	-0.009	-0.118
Guatemala	-0.096	-0.103	0.007	-0.044	-0.055	0.011	-0.053	-0.015	-0.038
Haiti	-0.300	-0.123	-0.177	-0.117	-0.110	-0.007	-0.184	0.034	-0.217
Honduras	-0.073	-0.038	-0.036	-0.073	-0.089	0.016	-0.001	0.003	-0.003
Jamaica	-0.037	-0.078	0.041	-0.022	-0.075	0.053	-0.015	0.021	-0.037
Nicaragua	-0.055	0.010	-0.065	-0.086	-0.051	-0.035	0.031	0.043	-0.012
Panama	0.050	0.082	-0.032	0.048	0.052	-0.004	0.003	0.035	-0.032
Paraguay	-0.203	-0.093	-0.110	-0.096	-0.121	0.025	-0.107	0.027	-0.133
Peru	-0.003	0.044	-0.047	0.003	-0.009	0.012	-0.006	0.053	-0.059
Trinidad & Tobago	0.046	0.023	0.022	0.036	0.072	-0.036	0.009	0.010	-0.001
Uruguay	0.040	0.038	0.001	0.061	0.063	-0.002	-0.021	0.026	-0.047
Venezuela	-0.061	0.024	-0.085	0.013	0.026	-0.014	-0.074	0.060	-0.134
Average	-0.065	-0.021	-0.044	-0.024	-0.029	0.005	-0.042	0.030	-0.071

Notes: This table provides the estimation of the conditional distribution of output based on 10 years averages of population growth, human capital stocks and investment rates for the Latin American countries. This is done for the years 1960 and 2000 and then the measure of conditional σ -convergence is constructed accordingly. The results are divided to overall effect, covariate effect, and time effect.

The fact that I have grouped countries by region to explore conditional σ -convergence is ad hoc. For each region there are also countries that have converged. For example, in Sub-Saharan Africa Kenya, Lesotho, Malawi, Mali, Mauritius, Niger, Senegal and Zimbabwe have converged. Of the Latin American countries, Argentina, Chile, El Salvador, Honduras, Panama, and Uruguay have converged while for the Asian countries, Afghanistan, India, Indonesia, Malaysia, Sri Lanka and Taiwan have converged. For MENA, Bahrain, Cyprus and Malta have conditionally converged.

Former results show how the great bulk of after 1980 increase in variation is driven by African and Latin American countries and that time and covariate effects vary substantially among these groups of countries. This way, the additional step is an assessment of

Table 4.6: Overall, Time and Covariate Effects for Conditional Variance for Asian Countries

	1960-2000			1960-1980			1980-2000		
Country	Overall	Covariate	Time	Overall	Covariate	Time	Overall	Covariate	Time
Bangladesh	-0.090	-0.052	-0.038	-0.073	-0.087	0.013	-0.017	-0.064	0.047
China	-0.069	-0.086	0.017	-0.180	-0.041	-0.139	0.111	0.129	-0.018
Fiji	0.085	0.043	0.042	-0.071	0.074	-0.145	0.156	0.162	-0.006
Hong Kong	-0.157	-0.104	-0.054	-0.142	-0.012	-0.130	-0.015	0.026	-0.041
India	-0.080	-0.054	-0.026	-0.123	-0.081	-0.042	0.042	-0.023	0.065
Indonesia	-0.088	-0.061	-0.027	-0.185	-0.108	-0.077	0.097	0.037	0.060
Malaysia	-0.007	0.061	-0.069	-0.048	0.063	-0.111	0.041	0.019	0.022
Nepal	-0.067	-0.057	-0.009	-0.038	-0.128	0.090	-0.028	-0.102	0.073
Pakistan	0.057	0.115	-0.058	0.013	0.068	-0.055	0.044	0.021	0.023
Papua Guinea	-0.066	-0.062	-0.004	-0.055	-0.146	0.091	-0.011	-0.104	0.093
Philippines	-0.004	0.038	-0.042	-0.046	0.043	-0.089	0.042	0.012	0.029
Singapore	-0.122	-0.009	-0.113	-0.219	-0.040	-0.178	0.097	0.040	0.057
South Korea	-0.019	-0.153	0.134	-0.135	-0.119	-0.017	0.116	0.125	-0.009
Sri Lanka	0.015	-0.015	0.030	-0.096	0.017	-0.113	0.111	0.078	0.033
Taiwan	-0.106	-0.148	0.042	-0.097	-0.034	-0.063	-0.010	0.058	-0.067
Thailand	0.028	0.093	-0.065	0.023	0.047	-0.024	0.005	-0.101	0.106
Average	-0.043	-0.028	-0.015	-0.092	-0.030	-0.062	0.049	0.020	0.029

Notes: This table provides the estimation of the conditional distribution of output based on 10 years averages of population growth, human capital stocks and investment rates for the Asian countries. This is done for the years 1960 and 2000 and then the measure of conditional σ -convergence is constructed accordingly. The results are divided to overall effect, covariate effect, and time effect.

each variable by isolating its variation from the others. I do this, by implementing the decomposition by holding all the variables fixed except one over time⁹. This strategy allows me to answer the question: ‘What would it happen if we only observed changes in education over time?’. So I may get insights about the leading determinant of the observed divergence process for each country and time period. Results are shown in Table 4.8 below.

Broadly speaking, the biggest effects are in the human capital decomposition, so that a variance that is conditional only to changes in the years of education is higher with respect to the other two variables. The impact is more relevant in the second period and especially for the two diverging areas. Once again I find heterogeneity over variables, time and regions. For instance I may see as population growth has a bigger divergent impact for Latin American countries with respect to the other covariates, while for Asian countries the investment rates over the last 20 years push towards a greater reduction of the variance. It could give insights about the capital movements towards the Asian countries in those decades, that may have favoured a shift in production function, given the inverted time effects I observe

⁹This in the spirit of decompositions as Machado & Mata (2005).

Table 4.7: Overall, Time and Covariate Effects for Conditional Variance for MENA Countries

	1960-2000			1960-1980			1980-2000		
Country	Overall	Covariate	Time	Overall	Covariate	Time	Overall	Covariate	Time
Algeria	-0.056	-0.090	0.034	-0.017	-0.049	0.032	-0.039	-0.033	-0.006
Cyprus	-0.051	-0.147	0.095	0.008	-0.102	0.110	-0.059	-0.025	-0.034
Egypt	-0.053	-0.034	-0.020	-0.090	-0.047	-0.043	0.037	-0.134	0.171
Iran	0.017	-0.068	0.084	0.010	0.007	0.003	0.006	0.003	0.003
Israel	-0.019	-0.009	-0.011	-0.010	-0.033	0.022	-0.009	0.006	-0.015
Jordan	-0.002	0.002	-0.005	-0.056	-0.058	0.002	0.053	0.030	0.023
Romania	-0.024	-0.083	0.059	0.012	-0.111	0.123	-0.037	0.015	-0.051
Syria	-0.044	-0.014	-0.029	-0.052	-0.032	-0.020	0.009	-0.097	0.106
Average	-0.029	-0.055	0.026	-0.024	-0.053	0.029	-0.005	-0.029	0.025

Notes: This table provides the estimation of the conditional distribution of output based on 10 years averages of population growth, human capital stocks and investment rates for the Middle East & North African (MENA) countries. This is done for the years 1960 and 2000 and then the measure of conditional σ -convergence is constructed accordingly. The results are divided to overall effect, covariate effect, and time effect.

in the subperiods.

4.5 Implications and Future Research

This paper was aimed to take a step towards the systematic study of the changes in variance of the conditional distribution of cross-country per capita income over time. While conditional β -convergence is likely to dominate empirical discussions of growth, the method presented here will make the study of distributional variation over time more appealing to growth empiricists. This can be used to determine if factor accumulation will indeed diminish the apparent increase in variation of output levels over time that has been at the crux of many growth debates. My results show that the great bulk of divergence among groups is given by Latin American and African countries during last twenty years, while Asian countries had been converging in the same period and OECD had the same behaviour until 1980 and then a stable distribution. By looking to the specific effects, changes in growth determinants covariates worked in order to increase the conditional variance for the period 1960-1980, while these effects were reverted after 1980 for almost all of the countries, unless Africa that is by far, the most divergent group in the second period. On the other hand the strong negative time effects in the period 1980-2000 suggest that is mainly something different from production factors accumulation that expanded output dispersion, where the main candidates are TFP and production factors returns. My feeling is that the time dimension needs to be explored more heavily and some omitted growth determinants such as institutions, trade policy, financial development, democracy, ethnic fractionalization and the like could all contribute to the time effect dominating the covariate effect.

Table 4.8: Conditional Variance Effects by Variable and Region

	Education			Population			Investment		
Country	Overall	Covariate	Time	Overall	Covariate	Time	Overall	Covariate	Time
1960-00									
Average	-0.072	0.003	-0.074	-0.061	0.001	-0.061	-0.058	-0.006	-0.053
Av. OECD	0.024	-0.056	0.080	0.019	-0.016	0.035	0.023	-0.004	0.027
Av. LAC	-0.085	-0.003	-0.082	-0.104	0.009	-0.113	-0.079	-0.006	-0.072
Av. ASIA	-0.049	0.024	-0.073	-0.048	0.000	-0.048	-0.039	-0.009	-0.030
Av. AFRICA	-0.164	0.055	-0.219	-0.111	0.008	-0.119	-0.126	-0.003	-0.123
Av. MENA	-0.038	-0.040	0.002	-0.029	0.003	-0.032	-0.043	-0.012	-0.031
1960-80									
Average	-0.019	-0.008	-0.011	0.002	-0.001	0.003	-0.003	-0.008	0.004
Av. OECD	0.011	-0.045	0.057	0.019	-0.014	0.033	0.009	-0.001	0.010
Av. LAC	-0.032	-0.009	-0.023	-0.023	0.009	-0.032	-0.033	-0.011	-0.022
Av. ASIA	-0.070	0.001	-0.071	-0.048	-0.006	-0.042	-0.051	-0.010	-0.041
Av. AFRICA	-0.002	0.025	-0.027	0.042	0.010	0.033	0.038	-0.006	0.044
Av. MENA	-0.040	-0.027	-0.013	-0.027	-0.011	-0.016	-0.018	-0.023	0.005
1980-00									
Average	-0.054	-0.027	-0.027	-0.043	0.001	-0.044	-0.041	-0.004	-0.036
Av. OECD	-0.001	0.005	-0.005	-0.006	0.007	-0.013	-0.002	-0.004	0.002
Av. LAC	-0.061	0.020	-0.082	-0.080	0.008	-0.088	-0.055	0.002	-0.057
Av. ASIA	0.043	0.011	0.032	0.044	-0.010	0.054	0.053	0.006	0.047
Av. AFRICA	-0.158	-0.119	-0.040	-0.106	-0.001	-0.105	-0.120	-0.013	-0.107
Av. MENA	-0.014	0.006	-0.019	-0.005	-0.013	0.009	-0.019	-0.012	-0.007

Notes: This table provides the estimation of the conditional distribution of output by isolating each variable's variation from the others. The decomposition is implemented by holding all the variables fixed except one over time. This is done for the years 1960 and 2000 and then the measure of conditional σ -convergence is constructed accordingly. The results are divided to overall effect, covariate effect, and time effect.

Chapter 5

Conclusions, Contributions, and Future Work

This research tackles the issue of model uncertainty in empirical growth studies as well as model uncertainty in empirical well-being and happiness literature. First, it highlights the major criticisms and limitations of cross-country growth regressions and happiness models, theory uncertainty, parameter uncertainty, and functional form uncertainty. It then aims to amplify the existing statistical techniques that are being used in the literature to overcome the issue of model uncertainty in this scope of regressions. Specifically, this research uses recently developed model averaging methods to deal with theory uncertainty. These methods empower empiricists and econometricians to consider many explanatory variables simultaneously and identify the true effect of each single regressor on growth rate and also life satisfaction. To incorporate theory uncertainty into these regressions, a sound well-established statistical tool, Bayesian model averaging (BMA), has gained increasing attention.

The idea is to average over a variety of models via attaching prior probabilities to the models as well as to the parameters. Posterior probabilities for different models are then calculated based on data. In the next step, distributions for the parameters are calculated by averaging the posterior distributions under each model, weighted by its posterior model probability. The BMA approach is an important development used to identify which effects appear to be systematic features of the growth data and happiness. Studies have shown that using BMA to average over all possible combinations of potential regressors usually leads to better out-of-sample predictions. Given the importance of BMA in uncovering the key determinants of economic growth as well as well-being models, I provide an application of this method in chapter 3. Moreover, appendix A of this research describes how the user can implement a Bayesian model averaging exercise in the available statistical software to avoid time costs that arise in coding BMA estimators.

Although BMA has started to attract growth empiricists and those who work in the field of life satisfaction, there are still many open questions related to this approach that are crucial in both growth and life satisfaction context. Open questions include the optimal choice of priors on the models and parameters, the performance of BMA when the true model is not in the chosen class of models, and finally convergence issue in the use of Markov Chain Monte Carlo (MCMC) algorithm to average over models with many regressors. Having said that, Frequentist model averaging (FMA) techniques are potential alternatives to BMA and are much less complex than Bayesian methods. Accordingly, another aim of this project is to provide a review on the development of the FMA approaches. Recently great progress has been made in this field and a variety of asymptotically optimal FMA estimators have been proposed by researchers. This project, via chapter 2, provides the application of the new estimators within Frequentist framework in growth empirics and compare the results

with those obtained from a BMA exercise. An attractive contribution of this research to the FMA approach is that it provides a solution to a known issue with the FMA approach that arises when the number of potential regressors is large. As the number of regressors increases, the size of model space over which the FMA approach is conducted explodes exponentially, making the implementation of FMA computationally infeasible. I use the insights of Magnus & Durbin (1999) and Magnus et al. (2010) to address this issue. They propose a weighted average least squares (WALS) estimator that greatly reduces the size of model space within a Bayesian framework. I extend their findings to Frequentist framework and show using proper transformations operationalize FMA and greatly reduce the required computing time for the FMA.

The FMA approach is a relatively new topic and further development is needed. For example, there is not yet a formal statistical theory stating how the standard errors should be computed within a Frequentist framework. Another open question is how the FMA approach can be applied to complex data, including truncated and censored data as well as data with measurement errors. Providing a generalization of the FMA approach to more complex models, such as transformation models in survival analysis, is also another open area of research.

This brings me to the final aim of this research, shedding more lights on the researches related to the shape of the distribution of per capita income via using nonparametric methods. Nonparametric approach is firmly establishing itself as an important statistical tool in empirical growth studies to deal with functional form uncertainty. Chapter 4 investigates the reasons behind the fact that the variance of the distribution of per capita income has been constantly increasing over the past decades. The method provided in this chapter is an important step toward formulating σ -convergence in growth literature.

In closing, I believe that more work needs to be done in both growth econometrics and life satisfaction/well-being econometrics to systematically uncover the key determinants of economic growth and key determinants of happiness across countries, potential nonlinearities, and heterogeneity in growth processes.

References

- Acemoglu, D. (2009), *Introduction to modern economic growth*, Princeton University Press.
- Aitchison, J. & Aitken, C. (1976), ‘Multivariate binary discrimination by the kernel method’, *Biometrika* **63**(3), 413–420.
- Akaike, H. (1973), ‘Information Theory and an Extension of the Maximum Likelihood principle, in second International Symposium on Information Theory’, *Akadémiai Kiadó, Budapest* **1**, 267–281.
- Amini, S. & Parmeter, C. (2011), ‘Bayesian model averaging in R’, *Journal of Economic and Social Measurement* **36**(4), 253–287.
- Anderson, G. (2004), ‘Making inferences about the polarization, welfare and poverty of nations: A study of 101 countries 1970–1995’, *Journal of Applied Econometrics* **19**(5), 537–550.
- Azariadis, C. & Stachurski, J. (2003), ‘A forward projection of the cross-country income distribution’, *Institute of Economic Research, Kyoto University, Discussion paper series* **570**.
- Barro, R. & Lee, J.-W. (2010), A new data set of educational attainment in the world, 1950–2010, Working Paper 15902, National Bureau of Economic Research.
URL: <http://www.nber.org/papers/w15902>
- Battisti, M. & Parmeter, C. (2010), Convergence tools and mixture analysis, Working Papers WP07-10, CeLEG Center for Labor and Economic Growth, University Luiss.
- Beaudry, P., Collard, F. & Green, D. (2005), ‘Changes in the world distribution of output per worker, 1960–1998: How a standard decomposition tells an unorthodox story’, *Review of Economics and Statistics* **87**(4), 741–753.
- Besag, J., Green, P., Higdon, D. & Mengersen, K. (1995), ‘Bayesian computation and stochastic systems (with discussion)’, *Statistical Science* **10**(1), 3–66.
- Bianchi, M. (1997), ‘Testing for convergence: evidence from non-parametric multimodality tests’, *Journal of Applied Econometrics* **12**(4), 393–409.
- Bjørnskov, C. (2003), ‘The Happy few: Cross-country evidence on social capital and life satisfaction’, *Kyklos* **56**(1), 3–16.
- Blanchflower, D. & Oswald, A. (2004), ‘Well-being over time in Britain and the USA’, *Journal of Public Economics* **88**(7), 1359–1386.
- Bliss, C. (2000), ‘Galton’s fallacy and economic convergence: a reply to Cannon and Duck’, *Oxford Economic Papers* **52**(1), 420–422.

- Bowman, A. (1984), 'An alternative method of cross-validation for the smoothing of density estimates', *Biometrika* **71**(2), 353–360.
- Brock, W. & Durlauf, S. (2001), 'What have we learned from a decade of empirical research on growth? Growth empirics and reality', *World Bank Economic Review* **15**(2), 229–272.
- Brock, W., Durlauf, S. & West, K. (2003), 'Policy evaluation in uncertain economic environments', *Brookings Papers on Economic Activity* **34**(1), 235–322.
- Bruni, L. & Porta, P. (2005), *Economics and happiness: Framing the analysis*, Oxford University Press, USA.
- Buckland, S. T., Burnham, K. P. & Augustin, N. H. (1997), 'Model selection: An integral part of inference', *Biometrics* **53**(2), 603–618.
- Burnham, K. & Anderson, D. (2002), *Model selection and multimodel inference: A practical information-theoretic approach*, second edn, Springer Verlag, New York.
- Cannon, E. & Duck, N. (2000), 'Galton's fallacy and economic convergence', *Oxford Economic Papers* **52**(2), 415–419.
- Canova, F. (2004), 'Testing for convergence clubs in income per capita: A predictive density approach', *International Economic Review* **45**(1), pp. 49–77.
- Carl-Johan Dalgaard, C.-J. & Vastrup, J. (2001), 'On the measurement of σ -convergence', *Economics Letters* **70**(2), 283–287.
- Casella, G. & George, E. (1992), 'Explaining the Gibbs sampler', *The American Statistician* **46**(3), 167–174.
- Caselli, F. & Feyrer, J. (2007), 'The Marginal product of capital', *The Quarterly Journal of Economics* **122**(2), 535–568.
- Chib, S. & Greenberg, E. (1995), 'Understanding the Metropolis-Hastings algorithm', *American Statistician* **49**(4), 327–335.
- Claeskens, G., Croux, C. & Kerckhoven, J. V. (2006), 'Variable selection for logistic regression using a Prediction-Focused Information Criterion', *Biometrics* **62**(4), 972–979.
- Claeskens, G. & Hjort, N. (2003), 'The Focused Information Criterion (with discussion)', *Journal of the American Statistical Association* **98**(464), 900–916.
- Claeskens, G. & Hjort, N. (2008), 'Minimizing average risk in regression models', *Econometric Theory* **24**(2), 493–527.
- Clark, A. & Oswald, A. (1994), 'Unhappiness and unemployment', *The Economic Journal* **104**(424), 648–659.
- Cleveland, W. (1979), 'Robust locally weighted regression and smoothing scatterplots', *Journal of the American Statistical Association* **74**(368), 829–836.

- Clyde, M. (2010), *BAS: Bayesian Adaptive Sampling for Bayesian Model Averaging*. R package version 0.92.
URL: <http://cran.r-project.org/package=BAS>
- Clyde, M., Ghosh, J. & Littman, M. (2011), ‘Bayesian adaptive sampling for variable selection and model averaging’, *Journal of Computational and Graphical Statistics* **20**(1), 80–101.
- Danilov, D. & Magnus, J. (2004), ‘On the harm that ignoring pretesting can cause’, *Journal of Econometrics* **122**(1), 27–46.
- Davis, J. (1984), ‘New money, and old man/lady and ‘Two’s company’: Subjective welfare in the NORC general social surveys, 1972–1982’, *Social Indicators Research* **15**(4), 319–350.
- Desdoigts, A. (1999), ‘Patterns of economic development and the formation of clubs’, *Journal of Economic Growth* **4**, 305–330.
- Di Tella, R., MacCulloch, R. & Oswald, A. (2001), ‘Preferences over inflation and unemployment: Evidence from surveys of happiness’, *The American Economic Review* **91**(1), 335–341.
- Diener, E., Gohm, C., Suh, E. & Oishi, S. (2000), ‘Similarity of the relations between marital status and subjective well-being across cultures’, *Journal of Cross-Cultural Psychology* **31**(4), 419–436.
- Dolan, P., Peasgood, T. & White, M. (2008), ‘Do we really know what makes us happy? A review of the economic literature on the factors associated with subjective well-being’, *Journal of Economic Psychology* **29**(1), 94–122.
- Doppelhofer, G. & Weeks, M. (2009), ‘Jointness of growth determinants’, *Journal of Applied Econometrics* **24**(2), 209–244.
- Durlauf, S. & Johnson, P. (1995), ‘Multiple regimes and cross-country growth behaviour’, *Journal of Applied Econometrics* **10**(4), 365–384.
- Durlauf, S., Johnson, P. & Temple, J. (2005), Growth Econometrics, in P. Aghion & D. S.N., eds, ‘Handbook of Economic Growth’, Vol. 1 of *Handbook of Economic Growth*, Elsevier, chapter 8, pp. 555–677.
- Durlauf, S., Kourtellos, A. & Minkin, A. (2001), ‘The local Solow growth model’, *European Economic Review* **45**(4-6), 928–940.
- Durlauf, S. N., Kourtellos, A. & Tan, C. M. (2008), ‘Are any growth theories robust?’, *Economic Journal* **118**(527), 329–346.
- Easterlin, R. (2001), ‘Income and happiness: Towards an unified theory’, *The economic journal* **111**(473), 465–484.
- Egger, P. & Pfaffermayr, M. (2009), ‘On testing conditional sigma-convergence’, *Oxford Bulletin of Economics and Statistics* **71**(4), 453–473.

- Ehrlich, I. (1973), 'Participation in illegitimate activities: A theoretical and empirical investigation', *The Journal of Political Economy* **81**(3), 521–565.
- Eicher, T. S., Papageorgiou, C. & Raftery, A. E. (2011), 'Default priors and predictive performance in Bayesian model averaging with application to growth determinants', *Journal of Applied Econometrics* **26**(1), 30–55.
- Ellison, C. (1991), 'Religious involvement and subjective well-being', *Journal of Health and Social Behavior* pp. 80–99.
- Feldkircher, M. & Zeugner, S. (2009), Benchmark priors revisited: On adaptive shrinkage and the supermodel effect in Bayesian model averaging, IMF Working Papers 09/202, International Monetary Fund.
URL: <http://ideas.repec.org/p/imf/imfwpa/09-202.html>
- Fernandez, C., Ley, E. & Steel, M. (2001a), 'Benchmark priors for Bayesian model averaging', *Journal of Econometrics* **100**(2), 381–427.
- Fernandez, C., Ley, E. & Steel, M. (2001b), 'Model uncertainty in cross-country growth regressions', *Journal of Applied Econometrics* **16**(5), 563–576.
- Ferrer-i-Carbonell, A. & Frijters, P. (2004), 'How important is methodology for the estimates of the determinants of happiness?', *The Economic Journal* **114**(497), 641–659.
- Fiaschi, D. & Lavezzi, A. (2003), 'Distribution dynamics and nonlinear growth', *Journal of Economic Growth* **8**(4), 379–401.
- Foster, D. & George, E. (1994), 'The risk inflation criterion for multiple regression', *The Annals of Statistics* **22**(4), 1947–1975.
- Frey, B. & Stutzer, A. (2001), *Happiness and economics: How the economy and institutions affect human well-being*, Princeton University Press.
- Frey, B. & Stutzer, A. (2002), 'What can economists learn from happiness research?', *Journal of Economic Literature* **40**(2), 402–435.
- Furnival, G. & Wilson, R. (1974), 'Regressions by leaps and bounds', *Technometrics* **16**(4), 499–511.
- Gelfand, A. (1990), 'Sampling-based approaches to calculating marginal densities', *Journal of the American Statistical Association* **85**(410), 398–409.
- Geman, S. & Geman, D. (1984), 'Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images', *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**(6), 721–741.
- George, E. (1999), 'Discussion of "Model averaging and model search strategies" by M. Clyde', *Bayesian Statistics* **6**, 157–185.
- George, E. & Foster, D. (2000), 'Calibration and empirical Bayes variable selection', *Biometrika* **87**(4), 731–747.

- Gisbert, F. (2003), 'Weighted samples, kernel density estimators and convergence', *Empirical Economics* **28**(2), 335–351.
- Green, P. (1995), 'Reversible jump Markov Chain Monte Carlo computation and Bayesian model determination', *Biometrika* **82**(4), 711–732.
- Hall, P., Li, Q. & Racine, J. (2007), 'Nonparametric estimation of regression functions in the presence of irrelevant regressors', *The Review of Economics and Statistics* **89**(4), 784–789.
- Hall, P., Racine, J. & Li, Q. (2004), 'Cross-validation and the estimation of conditional probability densities', *Journal of the American Statistical Association* **99**(468), 1015–1026.
- Hansen, B. (2007), 'Least squares model averaging', *Econometrica* **75**(4), 1175–1189.
- Hansen, B. & Racine, J. (2009), Jackknife model averaging, Working paper, Department of Economics, University of Wisconsin-Madison.
- Harrison, D. & Rubinfeld, D. (1978), 'Hedonic housing prices and the demand for clean air', *Journal of Environmental Economics and Management* **5**(1), 81–102.
- Hastings, W. (1970), 'Monte Carlo sampling methods using Markov chains and their applications', *Biometrika* **57**(1), 97–109.
- Helliwell, J. & Putnam, R. (2004), 'The social context of well-being', *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **359**(1449), 1435–1446.
- Henderson, D., Parmeter, C. & Russell, R. (2008), 'Modes, weighted modes, and calibrated modes: evidence of clustering using modality tests', *Journal of Applied Econometrics* **23**(5), 607–638.
- Heston, A., Summers, R. & Aten, B. (2009), 'Penn World Table version 6.3.', *Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania*.
- Hjort, L. & Claeskens, G. (2003), 'Frequentist model average estimators', *Journal of the American Statistical Association* **98**(464), 879–899.
- Hoeting, J., Madigan, D., Raftery, A. & Volinsky, C. (1999), 'Bayesian model averaging: A tutorial', *Statistical Science* **14**(4), 382–417. With comments by Clyde, M. and Draper, D. and George, E.I. and a rejoinder by the authors.
- Huber, P. & Pfaffermayr, M. (2010), 'Testing for Conditional convergence in variance and skewness: The firm size distribution revisited', *Oxford Bulletin of Economics and Statistics* **72**(5), 648–668.
- Hughes, M. & Thomas, M. (1998), 'The continuing significance of race revisited: A study of race, class, and quality of life in America, 1972 to 1996', *American Sociological Review* pp. 785–795.

- Hyndman, R., Bashtannyk, D. & Grunwald, G. (1996), ‘Estimating and visualizing conditional densities’, *Journal of Computational and Graphical Statistics* **5**(4), 315–336.
- Johnson, P. (2000), ‘A nonparametric analysis of income convergence across the US states’, *Economics Letters* **69**(2), 219–223.
- Johnson, P. (2005), ‘A continuous state space approach to “Convergence by Parts”’, *Economics Letters* **86**(3), 317–321.
- Jones, C. (1997), ‘Convergence revisited’, *Journal of economic Growth* **2**(2), 131–153.
- Kass, R. E. & Wasserman, L. (1995), ‘A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion’, *Journal of the American Statistical Association* pp. 928–934.
- Kohler, H., Behrman, J. & Skytthe, A. (2005), ‘Partner+children= happiness? The effects of partnerships and fertility on well-being’, *Population and Development Review* **31**(3), 407–445.
- Kourtellis, A. (2003), ‘Modeling parameter heterogeneity in cross-country growth regression models’, *Reproduced: University of Cyprus*.
- Leamer, E. (1978), *Specification searches: Ad hoc inference with nonexperimental data*, John Wiley & Sons, Inc.
- Leeb, H. & Pötscher, B. (2003), ‘The finite-sample distribution of post-model-selection estimators and uniform versus nonuniform approximations’, *Econometric Theory* **19**(1), 100–142.
- Leeb, H. & Pötscher, B. (2005), ‘Model selection and inference: facts and fiction’, *Econometric Theory* **21**(1), 21–59.
- Leung, G. & Barron, A. (2006), ‘Information theory and mixing least-squares regressions’, *IEEE Transactions on Information Theory* **52**(8), 3396–3410.
- Ley, E. & Steel, M. (2009a), ‘On the effect of prior assumptions in Bayesian model averaging with applications to growth regression’, *Journal of Applied Econometrics* **24**(4), 651–674.
- Ley, E. & Steel, M. F. J. (2009b), ‘Comments on jointness of growth determinants’, *Journal of Applied Econometrics* **24**(2), 248–251.
- Li, K. (1987), ‘Asymptotic optimality for C_P , C_L , cross-validation and generalized cross-validation: discrete index set’, *The Annals of Statistics* **15**(3), 958–975.
- Li, Q. & Racine, J. (2003), ‘Nonparametric estimation of distributions with categorical and continuous data’, *Journal of Multivariate Analysis* **86**(2), 266–292.
- Li, Q. & Racine, J. (2007), *Nonparametric econometrics: Theory and practice*, Princeton University Press Princeton, New Jersey.

- Liang, F., Paulo, R., Molina, G., Clyde, M. & Berger, J. (2008), 'Mixtures of g priors for Bayesian variable selection', *Journal of the American Statistical Association* **103**(481), 410–423.
- Liu, J. (2008), *Monte Carlo strategies in scientific computing*, Springer Verlag.
- Liu, J., Liang, F. & Wong, W. (2000), 'The multiple-try method and local optimization in Metropolis sampling', *Journal of the American Statistical Association* **95**(449), 121–134.
- Liu, Z. & Stengos, T. (1999), 'Non-linearities in cross-country growth regressions: A semi-parametric approach', *Journal of Applied Econometrics* **14**(5), 527–538.
- Maasoumi, E., Racine, J. & Stengos, T. (2007), 'Growth and convergence: A profile of distribution dynamics and mobility', *Journal of Econometrics* **136**(2), 483–508.
- Machado, J. & Mata, J. (2005), 'Counterfactual decomposition of changes in wage distributions using quantile regression', *Journal of Applied Econometrics* **20**(4), 445–465.
- Madigan, D. & Raftery, A. (1994), 'Model selection and accounting for model uncertainty in graphical models using Occam's window', *Journal of the American Statistical Association* **89**(428), 1535–1546.
- Madigan, D., York, J. & Allard, D. (1995), 'Bayesian graphical models for discrete data', *International Statistical Review/Revue Internationale de Statistique* pp. 215–232.
- Magnus, J. & Durbin, J. (1999), 'Estimation of regression coefficients of interest when other regression coefficients are of no interest', *Econometrica* **67**(3), 639–643.
- Magnus, J. R., Powell, O. & Prüfer, P. (2010), 'A comparison of two model averaging techniques with an application to growth empirics', *Journal of Econometrics* **154**(2), 139–153.
- Magnus, J., Wan, A. & Zhang, X. (2011), 'Weighted average least squares estimation with nonspherical disturbances and an application to the Hong Kong housing market', *Computational Statistics & Data Analysis* **55**(3), 1331–1341.
- Mallows, C. (1973), 'Some comments on C_p ', *Technometrics* **15**(4), 661–675.
- Mankiw, N., Romer, D. & Weil, D. (1992), 'A contribution to the empirics of economic growth', *The quarterly journal of economics* **107**(2), 407–437.
- Masanjala, W. & Papageorgiou, C. (2008), 'Rough and lonely road to prosperity: A reexamination of the sources of growth in Africa using Bayesian model averaging', *Journal of Applied Econometrics* **23**(5), 671–682.
- Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A. & Teller, E. (1953), 'Equation of state calculations by fast computing machines', *The Journal of Chemical Physics* **21**(6), 1087–1092.

- Millar, P. (2011), ‘BIC: Stata module to evaluate the statistical significance of variables in a model’, Statistical Software Components, Boston College Department of Economics. URL: <http://econpapers.repec.org/RePEc:boc:bocode:s449507>
- Nadaraya, E. (1964), ‘On estimating regression’, *Theory of Probability and Its Applications* **9**, 141–142.
- Neal, R. (2003), ‘Slice Sampling’, *The Annals of Statistics* **31**(3), 705–741.
- Paapaa, R. & Van Dijk, H. (1998), ‘Distribution and mobility of wealth of nations’, *European Economic Review* **42**(7), 1269–1293.
- Pagan, A. & Ullah, A. (1999), *Nonparametric econometrics*, Cambridge University Press.
- Pesaran, M., Schleicher, C. & Zaffaroni, P. (2009), ‘Model averaging in risk management with an application to futures markets’, *Journal of Empirical Finance* **16**(2), 280–305.
- Pittau, M. (2005), ‘Fitting regional income distributions in the European Union’, *Oxford Bulletin of Economics and Statistics* **67**(2), 135–161.
- Pittau, M. & Zelli, R. (2006), ‘Empirical evidence of income dynamics across EU regions’, *Journal of Applied Econometrics* **21**(5), 605–628.
- Pittau, M., Zelli, R. & Johnson, P. (2010), ‘Mixture models, convergence clubs, and polarization’, *Review of Income and Wealth* **56**(1), 102–122.
- Pritchett, L. (1997), ‘Divergence, big time’, *The Journal of Economic Perspectives* **11**(3), 3–17.
- Quah, D. (1993a), ‘Empirical cross-section dynamics and tests of the convergence hypothesis’, *European Economic Review* **37**, 426–434.
- Quah, D. (1993b), ‘Galton’s fallacy and tests of the convergence hypothesis’, *The Scandinavian Journal of Economics* **95**(4), 427–443.
- Quah, D. (1996a), ‘Convergence empirics across economies with (some) capital mobility’, *Journal of Economic Growth* **1**(1), 95–124.
- Quah, D. (1996b), ‘Empirics for economic growth and convergence’, *European economic review* **40**(6), 1353–1375.
- Quah, D. (1996c), ‘Twin peaks: growth and convergence in models of distribution dynamics’, *The Economic Journal* **106**(437), 1045–1055.
- R Development Core Team (2010), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL: <http://www.r-project.org>
- Raftery, A. (1995), ‘Bayesian model selection in social research’, *Sociological Methodology* **25**(1), 111–163.

- Raftery, A., Hoeting, J., Volinsky, C., Painter, I. & Yeung, K. Y. (2010), *BMA: Bayesian Model Averaging*. R package version 3.13.
URL: <http://cran.r-project.org/package=BMA>
- Raftery, A., Madigan, D. & Hoeting, J. (1997), ‘Bayesian model averaging for linear regression models’, *Journal of the American Statistical Association* **92**(437), 179–191.
- Raftery, A., Madigan, D. & Volinsky, C. (1996), ‘Accounting for model uncertainty in survival analysis improves predictive performance’, *Bayesian statistics* **5**, 323–349.
- Rudemo, M. (1982), ‘Empirical choice of histograms and kernel density estimators’, *Scandinavian Journal of Statistics* **9**(2), 65–78.
- Sala-i-Martin, X. (1997), ‘I just ran two million regressions’, *American Economic Review* **87**(2), 178–83.
- Sala-i-Martin, X., Doppelhofer, G. & Miller, R. (2004), ‘Determinants of long-term growth: A Bayesian averaging of classical estimates (BACE) approach’, *American Economic Review* **94**(4), 813–835.
- Schomaker, M., Wan, A. T. K. & Heumann, C. (2010), ‘Frequentist model averaging with missing observations’, *Computational Statistics and Data Analysis* **54**, 3336–3347.
- Shmotkin, D. (1990), ‘Subjective well-being as a function of age and gender: A multivariate look for differentiated trends’, *Social Indicators Research* **23**(3), 201–230.
- Smith, A. & Roberts, G. (1993), ‘Bayesian computation via the Gibbs sampler and related Markov Chain Monte Carlo methods’, *Journal of the Royal Statistical Society. Series B (Methodological)* **55**(1), 3–23.
- Stack, S. & Eshleman, J. (1998), ‘Marital status and happiness: A 17-nation study’, *Journal of Marriage and the Family* pp. 527–536.
- Stamatakis, D. & Petrakis, P. (2005), Human capital, growth and convergence traps: Implications from a cross-country analysis, Technical report, mimeo, Athens National and Kapodistrian University.
- Stone, C. (1977), ‘Consistent nonparametric regression’, *The Annals of Statistics* **5**(4), 595–620.
- Strawderman, W. (1971), ‘Proper Bayes minimax estimators of the multivariate normal mean’, *The Annals of Mathematical Statistics* **42**(1), 385–388.
- Tanner, M. & Wong, W. (1987), ‘The calculation of posterior distributions by data augmentation’, *Journal of the American Statistical Association* **82**(398), 528–540.
- Turlach, R. & Weingessel, A. (2007), *quadprog: Functions to solve Quadratic Programming Problems*. R package version 1.4-11.

- Venables, W. N. & Ripley, B. D. (2002), *Modern applied statistics with S*, fourth edn, Springer, New York. ISBN 0-387-95457-0.
URL: <http://www.stats.ox.ac.uk/pub/MASS4>
- Waite, L. (1995), ‘Does marriage matter?’, *Demography* **32**(4), 483–507.
- Wan, A. T. K., Zhang, X. & Zou, G. (2010), ‘Least squares model averaging by Mallows criterion’, *Journal of Econometrics* **156**, 277–283.
- Wang, H., Zhang, X. & Zou, G. (2009), ‘Frequentist model averaging estimation: A review’, *Journal of Systems Science and Complexity* **22**(4), 732–748.
- Wang, M. & Van Ryzin, J. (1981), ‘A class of smooth estimators for discrete distributions’, *Biometrika* **68**(1), 301.
- Watson, G. (1964), ‘Smooth regression analysis’, *Sankhyā: The Indian Journal of Statistics, Series A* **26**(4), 359–372.
- Whittaker, J. (1990), *Graphical models in applied multivariate statistics*, Wiley New York.
- Yang, Y. (2001), ‘Adaptive regression by mixing’, *Journal of the American Statistical Association* **96**(454), 574–588.
- Yuan, Z. & Yang, Y. (2005), ‘Combining linear regression models: When and how?’, *Journal of the American Statistical Association* **100**(472), 1202–1214.
- Zellner, A. (1986), ‘On assessing prior distributions and Bayesian regression analysis with g-prior distributions’, *Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti* **6**, 233–243.
- Zellner, A. & Siow, A. (1980), ‘Posterior odds ratios for selected regression hypotheses’, *Trabajos de Estadística y de Investigación Operativa* **31**(1), 585–603.

Appendix A

Bayesian Model Averaging in R

A.1 Abstract

Bayesian model averaging has increasingly witnessed applications across an array of empirical contexts. However, the dearth of available statistical software which allows one to engage in a model averaging exercise is limited. It is common for consumers of these methods to develop their own code, which has obvious appeal. However, canned statistical software can ameliorate one's own analysis if they are not intimately familiar with the nuances of computer coding. Moreover, many researchers would prefer user ready software to mitigate the inevitable time costs that arise when hard coding an econometric estimator. To that end, this paper describes the relative merits and attractiveness of several competing packages in the statistical environment R to implement a Bayesian model averaging exercise.

A.2 Introduction

Bayesian model averaging (BMA) is an empirical tool to deal with model uncertainty in various milieus of applied science. In general, BMA is employed when there exist a variety of models which may all be statistically reasonable but most likely result in different conclusions about the key questions of interest to the researcher. As Raftery (1995, pg. 113) notes “In this situation, the standard approach of selecting a single model and basing inference on it underestimates uncertainty about quantities of interest because it ignores uncertainty about model form.” Typically, though not always, BMA focuses on which regressors to include in the analysis. The allure of BMA is that one can quickly determine models, or more specifically sets of explanatory variables, which possess high likelihoods. By averaging across a large set of models one can determine those variables which are relevant to the data generating process for a given set of priors used in the analysis. Each model (a set of variables) receives a weight and the final estimates are constructed as a weighted average of the parameter estimates from each of the models. BMA includes all of the variables within the analysis, but shrinks the impact of certain variables towards zero through the model weights. These weights are the key feature for estimation via BMA and will depend upon a number of key features of the averaging exercise including the choice of prior specified.

The implementation of BMA, which was first proposed by Leamer (1978, Sections 4.4-4.6), for linear regression models is as follows. Consider a linear regression model with a constant term, β_0 , and k potential explanatory variables $x_1, x_2, \dots, x_k$,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon. \quad (\text{A.1})$$

Given the number of regressors, we will have 2^k different combinations of right hand side variables indexed by M_j for $j = 1, 2, 3, \dots, 2^k$. Once the model space has been constructed, the posterior distribution for any coefficient of interest, say β_h , given the data Δ is

$$Pr(\beta_h|\Delta) = \sum_{j:\beta_h \in M_j} Pr(\beta_h|M_j, \Delta) Pr(M_j|\Delta) \quad (\text{A.2})$$

BMA uses each model's posterior probability, $Pr(M_j|\Delta)$, as weights. The posterior model probability of M_j is the ratio of its marginal likelihood to the sum of marginal likelihoods over the entire model space and is given by

$$Pr(M_j|\Delta) = Pr(\Delta|M_j) \frac{Pr(M_j)}{Pr(\Delta)} = Pr(\Delta|M_j) \frac{Pr(M_j)}{\sum_{i=1}^{2^k} Pr(\Delta|M_i) Pr(M_i)} \quad (\text{A.3})$$

where

$$Pr(\Delta|M_j) = \int Pr(\Delta|\beta^j, M_j) Pr(\beta^j|M_j) d\beta^j \quad (\text{A.4})$$

and β^j is the vector of parameters from model M_j , $Pr(\beta^j|M_j)$ is a prior probability distribution assigned to the parameters of model M_j , and $Pr(M_j)$ is the prior probability that M_j is the true model. The estimated posterior means and standard deviations of

$\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k)$ are then constructed as

$$E[\hat{\beta}|\Delta] = \sum_{j=1}^{2^k} \hat{\beta} Pr(M_j|\Delta), \quad (\text{A.5})$$

$$V[\hat{\beta}|\Delta] = \sum_{j=1}^{2^k} (Var[\beta|\Delta, M_j] + \hat{\beta}^2) Pr(M_j|\Delta) - E[\beta|\Delta]^2. \quad (\text{A.6})$$

For further discussions on BMA, including its limitations and implementation, I refer the reader to the comprehensive review of Bayesian model averaging by Hoeting et al. (1999).

The following sections provide an overview of three currently available packages in the statistical computing language of R (R Development Core Team 2010) that can implement a BMA empirical exercise. The main features under the user's control for each of the packages, including the set of prior probabilities and model sampling algorithms as well as the plot diagnostics available to visualize the results, are described. Several detailed examples to compare the performance of these different packages are also provided along with functioning R code in an appendix.

To my knowledge, R is the only mainstream statistical platform which offers a suite of routines to conduct a BMA analysis. The availability of BMA routines in other statistical software is limited. Neither Gauss nor Stata possess built-in packages which allow the user to implement a genuine, linear regression BMA.^{1,2} Matlab, while lacking a comprehensive BMA toolbox,³ supplies users with the core functionality of the BMS package (discussed below) via installation of the *BMS toolbox for Matlab*⁴. Fortran users have access to a ready-to-use BMA toolbox stemming from Fernandez et al.'s (2001b) publicly available code. And finally, while SAS provides some functionality for implementing BMA it is incapable of handling a large-scale BMA analysis.

Beyond my review of the functionality of the three available packages, I contrast estimates and posterior inclusion probabilities across the three packages with a mock empirical example. This is done both with a set of covariates that allows for full enumeration of the model space as well as requiring the implementation of a model space search mechanism which is what truly distinguishes the three packages. The time performance of the three packages as both the sample size and the covariate space increase is also supplied. Finally, I examine whether these packages can replicate the results of recently published econometric research

¹Millar (2011) has recently published a Stata module that uses the Bayesian Information Criterion (BIC) for estimating the probability that a variable is a part of the final model. This module, available at <http://fmwww.bc.edu/repec/bocode/b/bic.ado>, calculates the BIC statistic for all possible combinations of the independent variables.

²Gauss users can find the code used in Sala-i-Martin et al. (2004) to implement the BACE technique at <http://www.nhh.no/Default.aspx?ID=3075>.

³Matlab's Econometrics Toolbox comes with a function called `bma_g` that provides very basic BMA functionality.

⁴This toolbox, developed by Feldkircher & Zeugner (2009), is not a standard Mathworks toolbox. It is essentially a Matlab function that provides an interface to use the BMS package from within Matlab without seeing R.

that employs BMA techniques. Overall, all three packages share relative advantages against their peers, yet I advocate for the **BMS** package given its versatility with user defined priors as well as the numerous options to customize one's BMA analysis.

A.3 Available Packages for R

A.3.1 The BMS Package

The **BMS**⁵ package (Feldkircher & Zeugner 2009) employs standard Bayesian normal-conjugate linear model as the base model and “Zellner’s g prior” as the choice of prior structures for the regression coefficients. Since the form of the hyperparameter g is crucial in BMA analyses, the **BMS** package sets g equal to the sample size, usually known as the unit information prior (UIP). **BMS** also provides alternative formulations regarding the choice of g . The main function in the **BMS** package to implement a BMA regression analysis is `bms()`.

A.3.1.1 Model Sampling

Since enumerating all potential variable combinations becomes infeasible quickly for a large number of covariates, the **BMS** package uses a Markov Chain Monte Carlo (MCMC) samplers to gather results on the most important part of the posterior distribution when more than 14 covariates exist. The MCMC sampler walks through the model space using the Metropolis-Hastings algorithm⁶, which works as follows: Suppose that the current model at step i is M_i with posterior model probability $p(M_i|y, X)$. The MCMC sampler for the **BMS** package randomly draws a candidate model and then moves to this model if its marginal likelihood is superior to the marginal likelihood of the current model. In this algorithm, the number of times each model is kept will converge to the distribution of posterior model probabilities $p(M_i|y, X)$. The **BMS** package offers two different MCMC samplers to look at models within the model space. These two methods differ in the way they propose candidate models. The first method is called the *birth-death* sampler (`mcmc=bd`). In this case, one of the potential regressors is randomly chosen; if the chosen variable is already in the current model M_i , then the candidate model M_j will have the same set of covariates as M_i but drop the chosen variable. If the chosen covariate is not contained in M_i , then the candidate model will contain all the variables from M_i plus the chosen covariate; hence the appearance (birth) or disappearance (death) of the chosen variable depends if it already appears in the model. The second approach is called the *reversible-jump* sampler (`mcmc=rev.jump`). This sampler draws a candidate model by the birth-death method with 50% probability and with 50% probability the candidate model randomly drops one covariate with respect to M_i and randomly adds one random variable from the potential covariates that were not included in model M_i .

The precision of any MCMC sampling mechanism depends on the number of draws the

⁵BMS is an acronym for Bayesian Model Selection.

⁶See Metropolis et al. (1953), Hastings (1970), Chib & Greenberg (1995), and Liu (2008).

procedure runs through. Given that the MCMC algorithms used in the **BMS** package may begin using models which might not necessarily be classified as ‘good’ models, the first set of iterations do not usually draw models with high posterior model probabilities (PMP). This indicates that the sampler will only converge to spheres of models with the largest marginal likelihoods after some initial set of draws (known as the burn-in) from the candidate space. Therefore, this first set of iterations will be omitted from the computation of results. In the **BMS** package the argument (**burn**) specifies the number of burn-ins (models omitted), and the argument (**iter**) the number of subsequent iterations to be retained. The default number of burn-in draws for either MCMC sampler is 1000 and the default number of iteration draws (excluding burn-ins) 3000.

A.3.1.2 Model Priors

The **BMS** package offers considerable freedom in the choice of model prior. One can employ the uniform model prior as the choice of prior model size (**mprior**="uniform"), the Binomial model priors where the prior probability of a model is the product of inclusion and exclusion probabilities (**mprior**="fixed"), the Beta-Binomial model prior (**mprior**="random") that puts a hyperprior on the inclusion probabilities drawn from a Beta distribution, or a custom model size and prior inclusion probabilities (**mprior**="customk"). Of the three packages currently available in R this is the only package that allows for custom priors.

A.3.1.3 Alternative Zellner’s g Priors

Different mechanisms have been proposed in the literature for specifying g priors. The options in the **BMS** package are as follows

1. **g**="UIP"; Unit Information Prior (UIP), that corresponds to $g = N$, the sample size.⁷
2. **g**="RIC"; Sets $g = K^2$ and conforms to the risk inflation criterion.⁸
3. **g**="BRIC"; A mechanism that asymptotically converges to the unit information prior ($g = N$) or the risk inflation criterion ($g = K^2$). That is, the g prior is set to $g = \max(N, K^2)$.⁹
4. **g**="HQ"; Follows the Hannan-Quinn criterion asymptotically and sets $g = \log(N^3)$.
5. **g**="EBL"; Estimates a local empirical Bayes g -parameter.¹⁰
6. **g**="hyper"; Takes the “hyper- g ” prior distribution.¹¹

⁷See Fernandez et al. (2001a)

⁸See Foster & George (1994) for more details

⁹See Fernandez et al. (2001a)

¹⁰See George & Foster (2000) and Liang et al. (2008)

¹¹See Liang et al. (2008) and Feldkircher & Zeugner (2009)

A.3.1.4 Outputs

The main objects returned by `bms()` are the posterior inclusion probabilities (PIP) and posterior means and standard deviations. In addition, a call to this function returns a list of aggregate statistics including the number of draws, burn-ins, models visited, the top models, and the size of the model space. It also, returns the correlation between iteration counts and analytical PMPs for the best models¹², those with the highest posterior model probabilities (PMP).

A.3.1.5 Plot Diagnostics

BMS package users have access to plots of the prior and posterior model size distributions, a plot of posterior model probabilities based on the corresponding marginal likelihoods and MCMC frequencies for the best models that visualize how well the sampler has converged. A grid with signs and inclusion of coefficients vs. posterior model probabilities for the best models and plots of predictive densities for conditional forecasts are also produced.

A.3.2 The BAS Package

The BAS¹³ package (Clyde 2010) performs BMA in linear models using stochastic or deterministic sampling without replacement from posterior distributions. Prior distributions on coefficients are from Zellner's g prior or mixtures of g priors corresponding to the Zellner-Siow Cauchy Priors or the Liang hyper g priors (see Liang et al. 2008). Other model selection criteria include AIC and BIC. The main function in the BAS package to implement a BMA regression analysis is `bas.lm()`.

A.3.2.1 Model Sampling

If the number of covariates is less than 25, the BAS package enumerates all models, otherwise it implements three different search algorithms to find the models with the highest posterior probability. The first algorithm is named the Bayesian Adaptive Sampling algorithm¹⁴, (`method="BAS"`), that samples without replacement using random or deterministic sampling, (`random = "TRUE/FALSE"`). The Bayesian Adaptive Sampling algorithm samples models from the model space without replacement using the initial sampling probabilities, and will update the sampling probabilities using the estimated marginal inclusion probabilities. If the explanatory variables are orthogonal, then the deterministic sampler provides a list of the top models in order of their approximate posterior probability, and provides an effective search if the correlations of variables is small to modest. The second algorithm is the Adaptive MCMC sampler (`method="AMCMC"`). This sampling mechanism requires various

¹²Note that here and throughout the remainder of the document, the terminology “best models” or “best” corresponds to best as judged by the analytical likelihood.

¹³BAS is an acronym for Bayesian Adaptive Sampling.

¹⁴See Clyde, Ghosh & Littman (2011)

parameters to be tuned appropriately for the algorithm to converge reasonably well. The package currently handles these parameter tunings automatically by getting the computer to update tuning parameters and other choices during the course of sampling from the model space. The last method runs an initial MCMC to calculate marginal inclusion probabilities and then samples without replacement as in `method="BAS"` (`method="MCMC+BAS"`). The default number of models to sample is "NULL", meaning that a call to `bas.lm()` enumerates all combinations. The user can also indicate a model to initialize the sampling by passing a binary vector to the BAS functions (`bestmodel`). In default mode, sampling starts with the full model.

A.3.2.2 Model Priors

The family of prior distribution on the models is nearly identical to the BMS package allowing uniform, Bernoulli or Beta-Binomial distributions as priors (`modelprior="uniform"`, `"Bernoulli"`, `"beta.binomial"`).

A.3.2.3 Alternative g Priors

To determine the prior distribution for the regression coefficients, the user has the following choices:¹⁵

1. `g="AIC"`; Akaike information criterion.
2. `g="BIC"`; Bayesian information criterion.
3. `g="g-prior"`; Takes $g = N$, the sample size, corresponding to the UIP.
4. `g="ZS-null"`; Employs Zellner & Siow's (1980) suggestion. If the two models under comparison are nested, the Zellner-Siow strategy is to place a flat prior on common coefficients and a Cauchy prior on the remaining parameters. This option utilizes the null model as the base model and compares each model with the null model.
5. `g="ZS-full"`; The distinction between this method and `ZS-null` is the choice of base model. `ZS-full` method utilizes the full model (i.e., all covariates included) as the base model and compares all other models with the full model.
6. `g="hyper-g"`; This option uses a family of priors on g that provides improved mean square risk over ordinary maximum likelihood estimates in the normal means problem. Strawderman (1971) introduced this set of priors. An advantage of the hyper- g prior is that the posterior distribution of g given a model is available in closed form.
7. `g="hyper-g-laplace"`; This method is same as `hyper-g` and the only difference is that it uses a Laplace approximation to calculate the priors on g . This avoids issues with modes on the boundary and leads to improved solutions.

¹⁵See a thorough review of g priors used in BAS package in Liang et al. (2008).

8. `g="EB-local"`; This procedure estimates a separate g for each model. The local empirical Bayes (EB) estimate of g is the maximum marginal likelihood estimate constrained to be nonnegative.
9. `g="EB-global"`; The global EB technique sets one common g for all models and borrows strength from all models by estimating g from the marginal likelihood of the data, averaged over all models.

A.3.2.4 Outputs

The main objects returned from a call to `bas.lm()` are posterior inclusion probabilities, posterior mean and standard deviations for each coefficient, as in a call to `bms()`. Moreover, `bas.lm()` also returns the prior inclusion probability for each variable along with the posterior probability that each variable is non-zero.

A.3.2.5 Diagnostic Plots

The main plotting command in this package provides a panel of four graphs. The first is a plot of the residuals versus fitted values using the Bayesian model averaged coefficient estimates. The second is a plot of the cumulative marginal likelihoods of all the visited models. The third is a plot of log marginal likelihood vs. model dimension and finally the fourth plot shows the posterior marginal inclusion probabilities.

A.3.3 The BMA Package

The **BMA**¹⁶ package (Raftery et al. 2010) performs BMA analysis assuming a uniform distribution on model priors and using a simple BIC (Bayesian Information Criterion) approximation to construct the prior probabilities on the regressions coefficients. This package is built upon Raftery's (1995) algorithm. The main functions in the **BMA** package to implement a BMA regression analysis are `bicreg()` and `iBMA.bicreg()`.

A.3.3.1 Model Sampling

If the number of covariates is less than 30, the **BMA** package enumerates all models. **BMA** first uses a backward elimination procedure, in which variables are eliminated one at a time starting from the full model, to reduce the initial set of variables to 30. The package screens the t-statistics for the estimated parameters and removes the variables for which statistics are small. It then utilizes a specific BIC difference, according to Table A.1, to compare model M_j to M_i to find models that are more likely a part of the final set of good models. The models that are left are said to belong to *Occam's window*, a generalization of the famous Occam's razor, or principle of parsimony in scientific explanation. If the initial

¹⁶BMA is an acronym for Bayesian model averaging.

set of models in Occam's window is large and all models are linear, **BMA** uses the leaps and bounds algorithm of Furnival & Wilson (1974) to select a reduced set of good models. The user can set a number specifying the maximum ratio for excluding models in Occam's window (**OR**). The default value of this ratio is 20. The maximum number of columns in the design matrix (**maxCol**), including the intercept, to be kept defaults to 31.

Table A.1: Grades of Evidence Corresponding to Values of the BIC Difference for M_j Against M_i

BIC Difference	Evidence
0–2	Weak
2–6	Positive
6–10	Strong
>10	Very strong

A.3.3.2 Model Priors

The **BMA** package is inflexible regarding model priors, assuming all models are on equal footing *a priori*. The model prior is set to $1/2^K$ where K is the total number of variables included in the analysis.

A.3.3.3 Regression Coefficients Priors

Unlike the **BMS** and **BAS** packages, **BMA** does not employ Zellner's g priors as its choice of the prior distributions for the regression coefficients. Instead, the package employs the BIC approximation, which corresponds fairly closely to the UIP. Specifically, for any coefficient, say β_1 , **BMA** calculates the BIC for all models that include β_1 and then the posterior probability that β_1 is in the model would be

$$Pr(\beta_1 \neq 0|\Delta) = \sum_{j:\beta_1 \in M_j} p(M_j|\Delta), \text{ such that: } p(M_j|\Delta) = \exp(-BIC_j/2) / \sum_{i=1}^K \exp(-BIC_i/2). \quad (\text{A.7})$$

A.3.3.4 Outputs

The call to **bicreg()** returns the posterior inclusion probabilities, the posterior mean and standard deviation of each coefficient (as with **bms()** and **bas.lm()**), along with the values for the BIC of the models selected, the posterior mean of each coefficient conditional on the variable being included in the model, and the R^2 for the models deemed to be the most important.

A.3.3.5 Diagnostic Plots

BMA generates a plot of the posterior distribution for each of the estimated coefficients.

A.4 Comparison of the Packages

Having briefly described each of the **BMS**, **BAS** and **BMA** packages, I present the main function calls available to the user in Table A.2 for ease of reference. The remainder of this section culls similarities of the key features and provides a detailed comparison of the available options for model sampling and search, construction of model priors, call outputs and diagnostics so that the reader is presented with enough information to judge the merits of each package.

A.4.1 Model Sampling

When the covariate space is small all three packages enumerate the model space and estimate all models to conduct the averaging. However, when the sample space of covariates is large, each of the packages engages in totally different search behavior across the space of candidate models. Both the **BAS** and **BMS** packages use adaptive search algorithms to determine a feasible list of models to construct the posterior distribution of model probabilities. These give added credibility to these packages relative to the **BMA** package which uses a much simpler mechanism to adjudicate across candidate models. As we will see in my empirical examples, when I enumerate the model space both **BMS** and **BAS** provide roughly identical results (posterior inclusion probabilities, means and standard deviations), whereas when I resort to model sampling the interpretations of the effects and inclusion of specific variables can be different. The **BMA** package produces roughly similar estimates of posterior means and standard deviations, but deviates substantially with respect to the inclusion probabilities.

A.4.2 Model Priors

As one can see, **BAS** is comparable to **BMS** in alternative options for model priors except that the **BMS** package offers the opportunity to construct a prior outside of the standard options available, making it more versatile for specific applications when specification of the g-prior is paramount.

A.4.3 Alternative Zellner's g Priors

The **BAS** package provides additional options, including BIC and AIC, relative to the **BMS** package. Recall that the **BMA** package does not provide options for coefficient priors consistent with either of the other packages.

Table A.2: A Comparison of BMS, BAS, and BMA Options

Package	Option	Description	Default
BMS:	<code>mcmc</code>	Indicates the model sampler to be used. If <code>bd</code> , a birth/death MCMC algorithm will be used. If <code>rev.jump</code> , a reversible jump algorithm is used that adds a swap step to the birth/death algorithm. If <code>enumerate</code> , the entire model space will be enumerated	<code>bd</code>
	<code>burn</code>	The number of burn-in draws for the MCMC sampler. Not taken into account if <code>mcmc=enumerate</code>	1000
	<code>iter</code>	Shows the number of iteration draws to be sampled excluding burn-ins	3000
	<code>nmodel</code>	The number of best models for which information is stored	500
	<code>g</code>	Indicates the hyperparameters on Zellner's g -prior for the regression coefficients. If <code>UIP</code> , the number of observations, $g = N$, is used as g -prior. If <code>BRIC</code> , $g = \max(N, K^2)$, where K is the number of covariates, will be used. If <code>RIC</code> , $g = K^2$ is considered. If <code>HQ</code> , the package sets $g = \log(N)^3$. If <code>EBL</code> , a local empirical Bayes g -parameter will be used. If <code>hyper</code> , the package takes the hyper- g prior distribution	N
	<code>mprior</code>	Indicates the choice of model prior. If <code>fixed</code> , then fixed common prior inclusion probabilities for each regressor is employed. The <code>random</code> option employs the random theta prior. If <code>uniform</code> , the uniform model prior will be utilized. The <code>customk</code> option, allows for custom model size priors. Alternatively, the <code>pip</code> option allows for custom prior inclusion probabilities	random
	<code>start.value</code>	Specifies the starting model of the iteration algorithm	NA
	<code>g.stats</code>	If <code>TRUE</code> , statistics on the shrinkage factor $g/(1+g)$ will be collected. Set to <code>FALSE</code> for faster iteration	<code>TRUE</code>
	<code>force.full.ols</code>	If <code>TRUE</code> , the OLS estimation part of the sampling algorithm relies on slower matrix inversion procedure. Setting this option to <code>TRUE</code> , slows down sampling but deals better with highly collinear data	<code>FALSE</code>
	<code>fixed.reg</code>	Includes the explanatory variables that the user wants to be always included in every sampled model	<code>numeric(0)</code>
	BAS:		
	<code>n.models</code>	Specifies the number of models to sample. If <code>NULL</code> , this package enumerates the entire model space unless $K > 25$	<code>NULL</code>
BAS:	<code>prior</code>	Indicates the choice of prior distribution for regression coefficients. Choices include <code>AIC</code> , <code>BIC</code> , <code>g-prior</code> , <code>ZS-null</code> , <code>ZS-full</code> , <code>hyper-g</code> , <code>hyper-g-laplace</code> , <code>EB-local</code> , and <code>EB-global</code>	<code>ZS-null</code>
	<code>modelprior</code>	Shows the family of prior distribution on the models. Choices include <code>uniform</code> , <code>Bernoulli</code> , and <code>beta.binomial</code>	uniform
	<code>method</code>	Indicates the sampling method. If <code>BAS</code> , a Bayesian Adaptive Sampling without replacement will be used. The <code>MCMC+BAS</code> option runs an initial MCMC algorithm to calculate marginal inclusion probabilities and then samples without replacement. If <code>AMCMC</code> , an Adaptive, experimental MCMC is considered.	<code>BAS</code>
	<code>initprobs</code>	Introduces a vector of length K with the initial inclusion probabilities used for sampling without replacement	Uniform
	<code>prob.rw</code>	Indicates the probability of using the random-walk proposal	0.5
	<code>Burnin.iterations</code>	Specifies the number of iterations to discard when using any of the MCMC options	<code>NULL</code>
	<code>MCMC.iterations</code>	Provides the number of iterations to run MCMC algorithm	<code>NULL</code>
	<code>lambda</code>	A parameter used in the AMCMC algorithm	<code>NULL</code>
	<code>delta</code>	This is the truncation parameter that prevents sampling probabilities to degenerate to 0 or 1	0.025
	BMA:		
BMA:	<code>wt</code>	A vector of weights for regression	<code>rep(1,length(y))</code>
	<code>OR</code>	Indicates the maximum ratio for excluding models in Occam's window	20
	<code>strict</code>	If <code>FALSE</code> , all models whose posterior model probability is within a facto of $1/OR$ will be returned. If <code>TRUE</code> , a more parsimonious set of models where any model with a more likely submodel is eliminated	<code>FALSE</code>
	<code>maxCol</code>	Specifies the maximum number of columns in the design matrix to be kept	31
	<code>thresProbne0</code>	A number that indicates the probability threshold for including variables as a percent	5
	<code>nIter</code>	Determining the maximum number of iterations that should be run	100

Notes: This table presents the available options and their default settings for the current Bayesian model averaging packages, **BMS**, **BAS**, and **BMA** within the statistical software environment R. In addition, a brief description of each option is also provided. For more details, see sections A.3 and A.4.

A.4.4 Outputs

All three packages return the posterior probability of inclusion as well as the posterior mean and standard deviations for each coefficient. However, only the **BMS** package provides a suite of diagnostic information to help the user gauge if further adjustment of the BMA setup is needed or to likely robustness of the results to various options within the model averaging exercise.

A.4.5 Diagnostic Plots

All three packages provide diagnostic output for the user to understand the results of the model averaging exercise. However, both the **BMS** and **BAS** packages provide plots of the dimensionality of the model whereas the **BMA** package does not.

A.5 Empirical Illustration

This section compares the results of a basic BMA exercise using the **UScrime** dataset available in the **MASS** package (Venables & Ripley 2002) in R. This is a dataset on per capita crime rates in 47 U.S. states in 1960 (see Ehrlich 1973). There are 15 potential independent variables, all perceived to be associated with crime rates. The last two, probability of imprisonment and average time spent in state prisons, were the predictor variables of interest in Ehrlich's (1973) original study, while the other 13 were control variables. This small sample will allow me to compare the results from the three packages when I can enumerate the model space. However, as a likely difference in the performance of the three packages, and the results gleaned from them, will depend upon a large model space, I also consider the same dataset with 35 fictitious variables (generated as independent standard normal random deviates) included so that each package must engage in model sampling.

A.5.1 Enumeration of the Model Space

I first compare the results of performing basic BMA with the original **UScrime** dataset, which is small enough to fully enumerate the model space for all three of the separate BMA packages discussed above. Table A.3 shows the probabilities that each variable belongs to the final model (PIP). All three packages in this setup fully enumerate all $2^{15} = 32,768$ models. Column (i) shows the PIPs calculated by the **BMS** package. The call to this package uses a uniform distribution for model priors and unit information priors for the distribution of regressions coefficients. Column (ii) presents the PIPs computed by the **BAS** package. Model priors are set to have uniform distributions and the priors on the regressions coefficients come from Zellner-Siow's g priors. Column (iii) indicates the PIPs that the **BMA** package computes. The **BMS** and **BAS** packages produce nearly identical PIPs while those of **BMA** are markedly different. The different internal algorithm for calculating Bayes factors and other parameters is the main reason for this disparity. Table A.4, on the other hand, shows the

posterior means and standard errors of regressors calculated from the three packages. Similar to the posterior inclusion probabilities, the estimated coefficients and standard errors computed by the **BMS** and **BAS** packages are literally identical but the estimates obtained from the **BMA** package are relatively different from their counterparts.

Table A.3: A Comparison of Posterior Inclusion Probabilities (PIPs) Across the **BMS**, **BAS**, and **BMA** Packages. (Full Model Enumeration)

	Packages		
	BMS	BAS	BMA
	(i)	(ii)	(iii)
PIP:			
Income Inequality	0.99	0.99	1.00
Education	0.97	0.97	1.00
Imprisonment Probability	0.89	0.89	0.98
Males	0.85	0.85	0.97
Non-white	0.67	0.68	0.89
Police Expenditure 1960	0.66	0.66	0.74
Unemployment 35-39	0.59	0.60	0.83
Police Expenditure 1959	0.42	0.43	0.25
Time	0.33	0.36	0.44
State Population	0.33	0.35	0.28
GDP	0.31	0.35	0.33
Southern State Dummy	0.23	0.26	0.06
Unemployment 14-24	0.20	0.24	0.11
Labor Force	0.15	0.19	dropped

Notes: This table reports the posterior inclusion probabilities (PIP) obtained from applying Bayesian model averaging to a dataset on per capita crime rate in 47 U.S. states, originally published in Ehrlich (1973), using the **BMS**, **BAS**, and **BMA** packages. The dependent variable is the rate of crimes in a particular category per head of population across 47 states for the year 1960. All models estimated contain an intercept. Column (i) shows the Bayesian posterior inclusion probabilities computed by the **BMS** package. Column (ii) shows the same probabilities reported by the **BAS** package. The posterior inclusion probabilities reported by the **BMA** package are presented in column (iii). Note since the number of explanatory variables in this case is 15, the model space is small enough that allows for full enumeration of model space.

A.5.2 Model Sampling

In this section I compare the results across the three packages when there is a large model space and model search must be undertaken to ascertain the best models. I create a large model space by adding 35 independent, fictitious variables obtained from a standard-normal distribution to the original **UScrime** dataset (for a total of 50 covariates). Tables A.5 and A.6 present the results. Using the available controls for each package I tried my best to set

Table A.4: A Comparison of Posterior Means and Standard Deviations Across the BMS, BAS, and BMA Packages. (Full Model Enumeration)

	Packages		
	BMS	BAS	BMA
	(i)	(ii)	(iii)
Posterior Mean & Std. Dev.			
Income Inequality	1.41 (0.35)	1.40 (0.36)	1.38 (0.33)
Education	1.90 (0.61)	1.86 (0.62)	2.12 (0.51)
Imprisonment Probability	-0.21 (0.11)	-0.21 (0.11)	-0.24 (0.10)
Males	1.16 (0.67)	1.14 (0.66)	1.40 (0.53)
Non-white	0.06 (0.05)	0.06 (0.05)	0.09 (0.04)
Police Expenditure 1960	0.62 (0.52)	0.61 (0.53)	0.66 (0.43)
Unemployment 35-39	0.20 (0.21)	0.20 (0.21)	0.27 (0.19)
Police Expenditure 1959	0.32 (0.51)	0.32 (0.51)	0.22 (0.40)
Time	-0.07 (0.15)	-0.08 (0.16)	-0.12 (0.17)
State Population	-0.02 (0.03)	-0.02 (0.03)	-0.01 (0.03)
GDP	0.18 (0.35)	0.19 (0.36)	0.19 (0.35)
Southern State Dummy	0.03 (0.08)	0.03 (0.09)	0.006 (0.03)
Unemployment 14-24	-0.02 (0.15)	-0.02 (0.17)	-0.03 (0.13)
Labor Force	0.04 (0.27)	0.05 (0.30)	dropped —

Notes: This table reports the posterior means and standard errors of regressors calculated from the **BMS**, **BAS**, and **BMA** packages. The Bayesian model averaging is applied to a dataset on per capita crime rate in 47 U.S. states, originally published in Ehrlich (1973). The dependent variable is the rate of crimes in a particular category per head of population across 47 states for the year 1960. All models estimated contain an intercept. Column (i) shows the posterior means and standard deviations computed by the **BMS** package. Column (ii) shows the same statistics reported by the **BAS** package. The posterior means and standard errors of regressors reported by the **BMA** package are presented in column (iii). Standard errors are in the parentheses. Note since the number of explanatory variables in this case is 15, the model space is small enough that allows for full enumeration of model space.

up the example so that the three packages are as identical as possible.

Table A.5: Posterior Inclusion Probabilities Across the BMS, BAS, and BAS Packages.
(Model Sampling Required)

	Packages		
	BMS	BAS	BMA
	(i)	(ii)	(iii)
Post Probabilities:			
Income Inequality	0.99	1.00	1.00
Males	0.99	1.00	0.19
Imprisonment Probability	0.97	0.99	1.00
Education	0.97	0.99	1.00
Unemployment 35-39	0.74	0.93	1.00
Time	0.73	0.97	dropped
Police Expenditure 1960	0.57	0.89	dropped
Police Expenditure 1959	0.54	0.22	1.00
Southern State Dummy	0.51	0.79	0.42
Unemployment 14-24	0.33	0.12	0.91
GDP	0.31	0.17	0.97
Males Per Females	0.25	0.08	0.86
Non-white	0.25	0.04	0.99
Labor Force	0.23	0.04	dropped

Notes: This table reports the posterior inclusion probabilities (PIP) obtained from applying Bayesian model averaging to a dataset on per capita crime rate in 47 U.S. states, originally published in Ehrlich (1973), using the BMS, BAS, and BMA packages. The dependent variable is the rate of crimes in a particular category per head of population across 47 states for the year 1960. All models estimated contain an intercept. Column (i) shows the Bayesian posterior inclusion probabilities computed by the BMS package. Column (ii) shows the same probabilities reported by the BAS package. The posterior inclusion probabilities reported by the BMA package are presented in column (iii). In this exercise, 35 independent, fictitious variables generated from a standard-normal distribution are added to the original dataset. Therefore, the model space is large enough that forces the packages to use a model space search mechanism.

Table A.5 shows the PIPS. Column (i) presents the probabilities obtained from the BMS package. This package employs the *birth-death* sampler using an MCMC search algorithm to find the posterior probabilities with a burn-in of 2000 models. Its choices for model priors and the distribution of the model coefficients are uniform distribution and unit information priors, respectively. Column (ii) displays the results from the BAS package. I choose a uniform distribution for the model prior, Zellner-Siow's g prior for the distributions of the coefficients, and an adaptive MCMC method for model sampling. Column (iii) shows the PIPs that the BMA package assigns to each variable. As discussed earlier, the BMA package takes an entirely different model sampling approach than either the BMS or BAS packages. It first reduces the initial set of variables using a backward elimination algorithm and then implements the iterated BMA method for variable selection. The package calls repeatedly to

a BMA procedure, iterating through the variables in a fixed order. After each call only those variables which have posterior probability greater than a specified threshold, controlled through `thresProbne0` in the call to `iBMA.bicreg()`, are kept and those variables whose posterior probabilities do not meet the threshold are replaced with the next set of variables. The call to the BMA package uses the BIC approximation as its choice for the distributions of model coefficients which is roughly identical to UIPs.

Table A.6 shows the estimated posterior means and standard deviations of the covariates. Moreover, how long each of the packages takes to run is shown in the last line of the table. The BMS and BAS packages that use similar internal search algorithms for model sampling generate roughly identical posterior means and standard deviations, and fairly identical posterior probability that each variable is non-zero. The BMA package, the fastest of the three, however, calculates significantly different posterior probabilities, means and standard deviations from what BMS and BAS produce. This difference lies in the fact that the BMA package relies on a hierarchical OLS t -statistic to obtain a handful of likely models, directly conflicting with the model sampling approaches in the BMS and BAS packages.

A.5.3 Plot Diagnostics

Plotting is an important tool that helps the user to visualize the shape of the posterior distributions of coefficients, assess the final model size, compare PIPs and look at model complexity. The BMS, BMA, and BAS packages all plot the marginal inclusion densities as well as images that show inclusion and exclusion of variables within models using separate colors. However, the BMS and BAS packages provide more graphical visualizations for the users. The following figures are some examples of the plots that these packages provide. These figures all correspond to the above example where I fictitiously incorporated 35 white noise covariates to force the packages to engage in searching over the model space.

Figures A.1 and A.2 are combined plots provided by the BMS package. The upper plot in each figure shows the prior and posterior distribution of model sizes and helps to illustrate to the user the impact of the model prior assumption on the estimation results. Consistent with the research of Ley & Steel (2009a) and Eicher et al. (2011), who stress the importance of model priors in applied work, the plots produced by the BMS package allow for visual clarification of choice of model prior on posterior results. For example, the upper plot of Figure A.1 assumes a uniform distribution for the model prior whereas the upper plot of Figure A.2 assumes a “fixed” common prior inclusion probability for each regressor as an alternative to a uniform prior.¹⁷ These plots allow the user to graphically compare across a range of model size priors to determine the impact on the posterior distribution. In the example here it seems that the model prior (either fixed or uniform) has little effect on the posterior distribution of model sizes.

The lower plot of both figures shows the analytical likelihood of the best 200 models and their MCMC frequencies/draws of these models from the MCMC sampler. If the sampler has converged, then the MCMC draws should conform to the analytical/exact likelihoods.

¹⁷See Sala-i-Martin et al. (2004) for more details.

Table A.6: A Comparison of Posterior Means and Standard Deviations Across the BMS, BMA, and BAS Packages. (Model Sampling Required)

	Packages		
	BMS	BAS	BMA
	(i)	(ii)	(iii)
Posterior Mean & Std. Dev.			
Income Inequality	1.58 (0.31)	1.49 (0.22)	2.55 (0.31)
Males	1.65 (0.41)	1.55 (0.32)	0.08 (0.23)
Imprisonment Probability	-0.30 (0.10)	-0.34 (0.06)	-0.26 (0.05)
Education	1.72 (0.59)	1.64 (0.38)	2.97 (0.48)
Unemployment 35-39	0.23 (0.17)	0.26 (0.11)	0.67 (0.20)
Time	-0.29 (0.22)	-0.36 (0.13)	dropped —
Police Expenditure 1960	0.51 (0.53)	0.81 (0.37)	dropped —
Police Expenditure 1959	0.49 (0.55)	0.14 (0.35)	0.83 (0.14)
Southern State Dummy	0.08 (0.11)	0.14 (0.09)	0.08 (0.12)
Unemployment 14-24	0.07 (0.20)	0.02 (0.11)	-0.54 (0.28)
GDP	0.13 (0.29)	0.06 (0.19)	0.90 (0.38)
Males Per Females	-0.32 (1.03)	-0.10 (0.44)	-2.54 (1.58)
Non-white	0.004 (0.029)	0.001 (0.009)	0.10 (0.03)
Labor Force	0.03 (0.44)	0.01 (0.11)	dropped —
Computation Time			
(Seconds)	64	16	12

Notes: This table reports the posterior means and standard errors of regressors calculated from the BMS, BAS, and BMA packages. The Bayesian model averaging is applied to a dataset on per capita crime rate in 47 U.S. states, originally published in Ehrlich (1973). The dependent variable is the rate of crimes in a particular category per head of population across 47 states for the year 1960. All models estimated contain an intercept. Column (i) shows the posterior means and standard deviations computed by the BMS package. Column (ii) shows the same statistics reported by the BAS package. The posterior means and standard errors of regressors reported by the BMA package are presented in column (iii). Standard errors are in the parentheses. In this exercise, 35 independent, fictitious variables generated from a standard-normal distribution are added to the original dataset. Therefore, the model space is large enough that forces the packages to use a model space search mechanism. The bottom part of table shows the computation time in seconds.

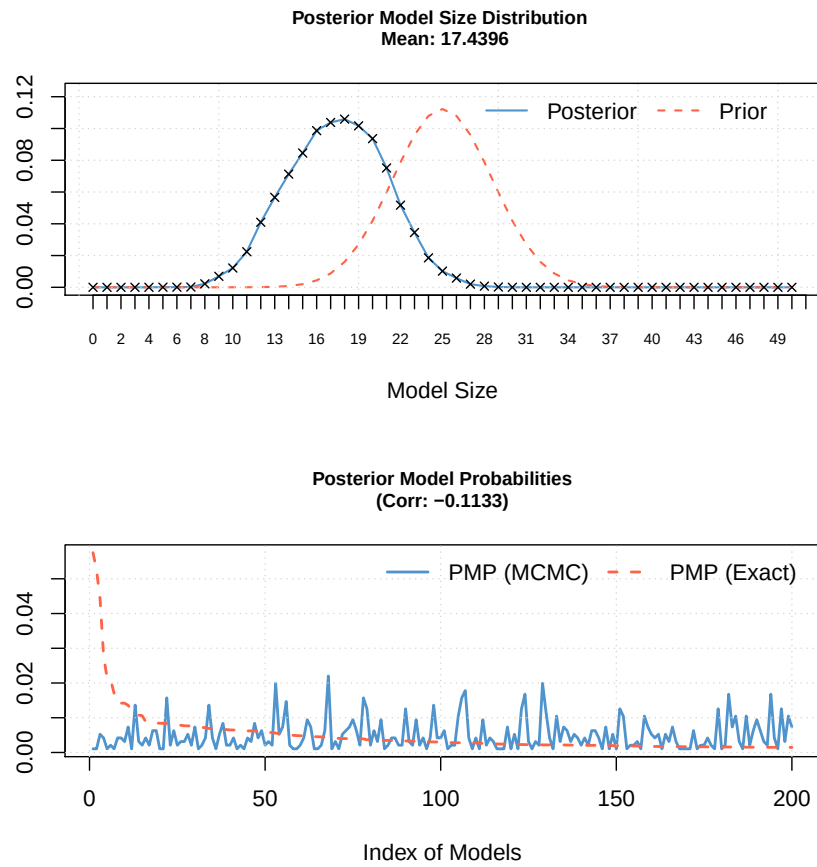


Figure A.1: Posterior Model Size Distribution and Model Probabilities Produced From the BMS Package With “Uniform” Model Priors.

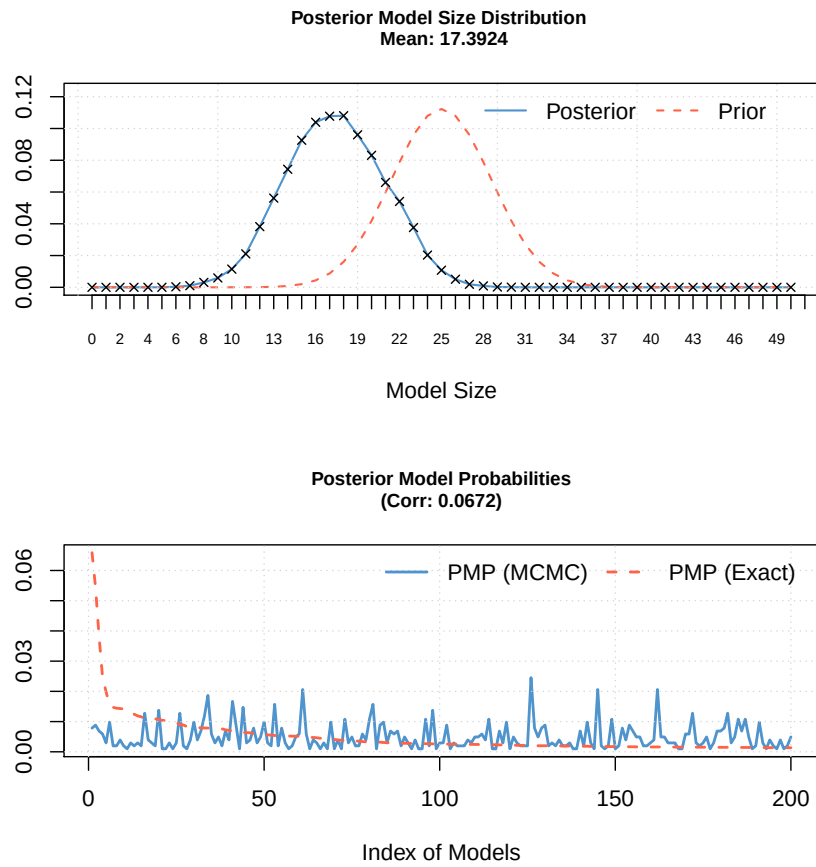


Figure A.2: Posterior Model Size Distribution and Model Probabilities Produced From the BMS Package With “Fixed” Model Priors.

This is expressed in the correlation of the two lines by the correlation reported in parentheses beneath the plot's title. In other words, these graphs are an indicator of how well the current 200 best models encountered by the search algorithm of the **BMS** package have converged. These plots are useful in determining the search mechanisms ability to find 'good' models throughout the model space. If the user deems that the best models are not acceptable then longer runs of the search algorithm or more burn-ins may be specified to help refine the search over the model space.

Figure A.3(a) is a visualization of the mixture marginal posterior density for the coefficient of GDP produced by the **BMS** package. The bar above the density shows the PIP, and the dotted vertical lines show the corresponding standard deviation bounds from the MCMC approach. Figure A.3(b) shows the same visualization using the **BMA** package. The posterior probability that the coefficient is zero is represented by a solid line at zero, with height equal to the probability. The nonzero part of the distribution is scaled so that the maximum height is equal to the probability that the coefficient is nonzero. Figure A.3(c) shows the same visualization using the **BAS** package. Similarly, the posterior probability that the coefficient is zero is represented by a solid line at zero, with height equal to the probability. The nonzero part of the distribution is scaled so that the maximum height is equal to the probability that the coefficient is nonzero. I mention here that Figure A.3(b) is the only plot produced within the **BMA** package.

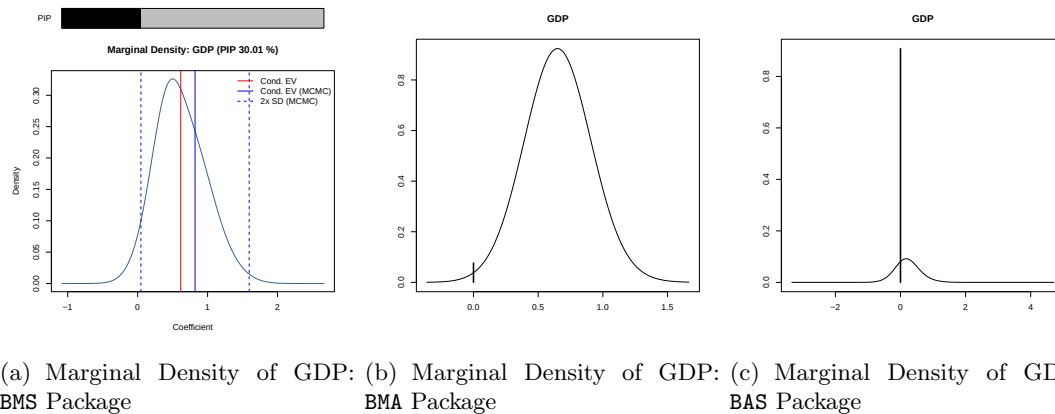


Figure A.3: Marginal Posterior Density Plots Produced Across the Three Packages

The plots of Figure A.4 are created within the **BAS** package. The first panel, (a), is a plot of the residuals versus the fitted values from the model averaging exercise. This graph allows the user to visualize if the fitted values are heavily dependent on the residuals. If one can see a specific pattern or trend in this graph, then it could be a sign that an important variable is missing. Panel (b) is a plot of the log marginal likelihood versus the overall model dimension. Using this graph, the user can visually predict the dimensionality of models with the highest marginal likelihoods. For instance the majority of the best models is concentrated around 10 to 12. Panel (c) shows the posterior marginal inclusion probabilities for each variable.¹⁸ This diagnostic plot makes it immediately obvious that the additional white noise variables

¹⁸The variables with PIPs>0.5 are drawn in red for ease of identification for the user.

I included almost uniformly have very low PIPs.

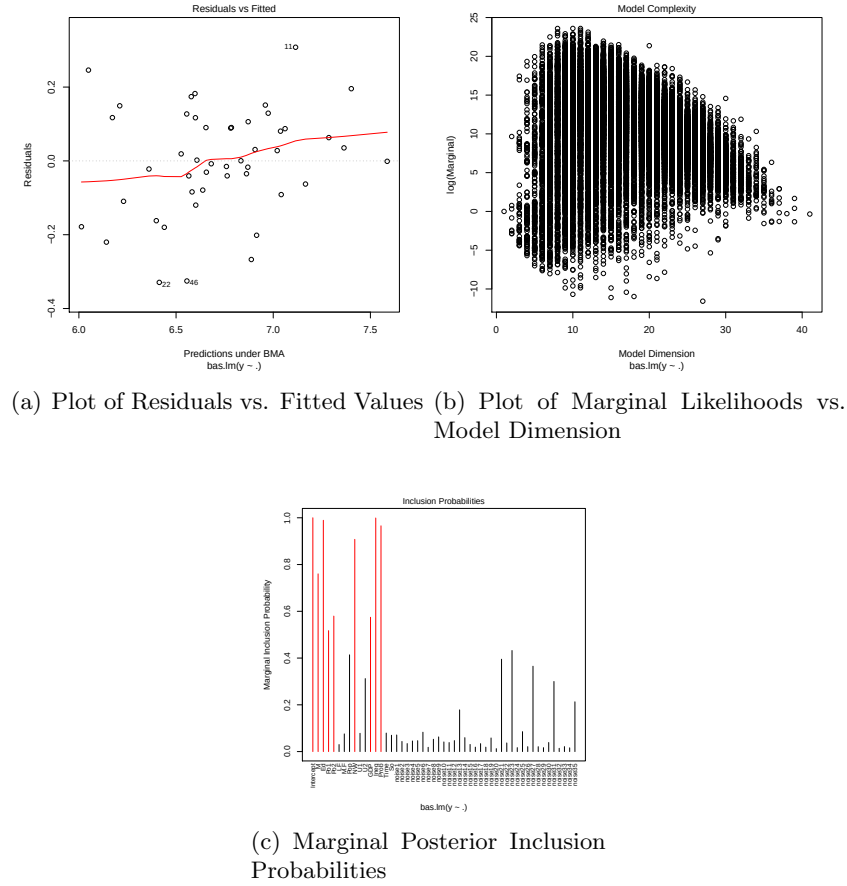


Figure A.4: Diagnostic Plots Produced With the BAS Package

A.6 Time Performance of the Three Packages

A key component of any statistical software is the amount of time it takes to produce results. To that end, this section reports the run times for the three packages across various sample sizes and dimensions of the covariate space. The size of the covariate space comes from $\{15, 30, 60, 120\}$ while the sample size is taken from $\{50, 500, 5000, 50000\}$. The dependent variable as well as the covariates are drawn from a standard normal distribution. I run my analysis on a 3.33GHz Intel(R) Core(TM) 2 Duo processor with 3.00 GB of DDR3 memory. Table A.7 presents the run times in seconds for each package. Within each set of parenthesis are the specific run times for a given covariate-sample size combination from baseline¹⁹ calls to the BMS, BAS, and BMA packages, respectively.

¹⁹Baseline options include uniform model priors, 500,000 iterations of the chosen MCMC algorithm, and discarding the first 3000 models (burn-in).

The first thing to notice is that due to the fact that the **BMA** package uses a simple model selection algorithm, it is the fastest among the three packages when the number of covariates is less than 30. The **BAS** package that provides roughly the same options as the **BMS** package is slower than the **BMA** package but noticeably faster than the **BMS** package. Once the number of covariates is greater than 15 and the sample size is beyond 500, the **BAS** package is uniformly the fastest of the packages. One reason that the **BAS** package realizes faster computing times relative to the **BMS** package with similar controls is that the **BAS** package is an R wrapper for a C library. This enables the call to the **BAS** package to take advantage of the high speed compiled C code used in the package.

Note that all three packages display increases in computation time as the dimensionality of the covariates increases. The number of covariates (intuitively) leads to larger increases in computation time than does increases in the sample size. Note that there are at least two counter-acting effects as the sample size, n , increases. A larger n puts more posterior mass on the best models. Therefore I expect that the MCMC sampler will converge quicker. On the other hand, the larger n , the more ‘large’ models will be emphasized, which slows down the sampler. If these two effects offset each other then I expect little change in run time over n . This is apparent in the **BMS** package. Notice how the **BMS** package has run times which decrease for a fixed level of covariates as I increase n while both the **BAS** and **BMA** packages actually take longer for $n = 50,000$ relative to $n = 5,000$.

Table A.7: A Comparison of BMS, BAS, and BMA Run Times.

Number of Covariates:	15	30	60	120
Sample Size:				
50	(44.86, 1.90, 0.74)	(42.30, 4.43, 2.17)	(42.33, 12.26, 8.15)	(58.42, 105.83, 101.69)
500	(39.03, 1.65, 0.50)	(40.44, 3.64, 2.78)	(39.36, 8.08, 24.67)	(42.46, 18.31, 75.21)
5,000	(34.97, 1.66, 1.00)	(35.70, 2.43, 4.78)	(38.89, 5.32, 38.09)	(39.78, 9.92, 81.70)
50,000	(34.05, 2.36, 6.95)	(34.40, 4.11, 13.02)	(38.89, 8.25, 34.47)	(44.45, 22.12, 195.68)

Notes: This table presents the run times in seconds for the three packages, **BMS**, **BAS**, and **BMA**, across various sample sizes and dimensions of the covariate space. Tabular Values Are Total Run Time in Seconds for the **BMS**, **BAS** and **BMA** Packages, Respectively. The size of the covariate space comes from $\{15, 30, 60, 120\}$ while the sample size is taken from $\{50, 500, 5000, 50000\}$. The dependent variable as well as the covariates are drawn from a standard normal distribution. The analysis is run on a 3.33GHz Intel(R) Core(TM) 2 Duo processor with 3.00 GB of DDR3 memory. Within each set of parenthesis are the specific run times for a given covariate-sample size combination from baseline calls to the **BMS**, **BAS**, and **BMA** packages, respectively. Baseline options include uniform model priors, 500,000 iterations of the chosen MCMC algorithm, and discarding the first 3000 models (burn-in).

A.7 Replicability

In this section I examine the ability of these three packages to replicate the results of published work deploying BMA using handwritten code. Fernandez et al. (2001b) (FLS hereafter) use a cross section of 72 countries along with 41 potential growth determinants

for the period 1960 to 1992.²⁰ FLS apply BMA to find the key determinants of economic growth given the numerous plausible models that have emerged on the topic. I use the same dataset and deploy all three BMA packages, **BMS**, **BAS**, and **BMA**, to attempt to replicate their results. In order to maximize my opportunity to replicate the FLS results, I set the available options within each of the three packages as close as possible to the specifications listed in FLS. The **BMS** package applies the MCMC algorithm to search over the model space, burns the first 100,000 models and the number of iteration draws to be sampled by its MCMC sampler is 200,000. It assigns the uniform distribution to the model priors. Similarly, the **BAS** package employs MCMC method to walk through the model space, discards the first 100,000 models, draws samples from the model space 200,000 times, and sets the models priors to the uniform distribution. The **BMA** package, on the other hand, does not have enough options to directly mimic the setup in FLS (see sections A.3.3 and A.4.1). Having said this, I set the number of iteration draws used by its search algorithm to 200,000, the maximum ratio for excluding models in Occam's window, **OR**, to 20 and keep the the maximum number of columns in the design matrix at the default of 31.

Table A.8 shows the PIPs for the variables of interest. I do not present posterior means or standard deviations since FLS only reported the PIPs in the body of their paper. To eschew making statements regarding results which FLS did not cover, I focus exclusively on the ability of the packages to reproduce the PIPs found in FLS. Column (i) shows the published PIPs that Fernandez et al. (2001*b*) have reported in their work and the remainder of table presents the PIPs computed via the **BMS**, **BAS**, and **BMA** packages.

As is apparent, only the **BMS** package is reasonably successful at matching the reported PIPs in FLS while the PIPs produced by the **BAS** package display significant differences (compare PIPs marked by *). The **BMA** package also fails achieve the same PIPs of FLS.²¹ This most likely lies in the fact that the **BMA** package was not called using exactly the setup in FLS and the difference in searching the model space that was described earlier. Interestingly, the PIPs returned from the **BMS** package almost uniformly match FLS' PIPs greater than 0.5. A key distinction between the results is that both the **BAS** and **BMA** packages suggest a set of variables that belong in the final model ($PIP > 0.5$) beyond those found in FLS. Specifically, the **BAS** package finds 13 variables with PIPs > 0.5 beyond FLS (and one variable with $PIP < 0.5$ from FLS) while the **BMA** package finds 10 variables with PIPs > 0.5 (and one variable with $PIP < 0.5$ from FLS).

To further test the limits of these packages to replicate published results on BMA I attempt to reproduce the estimates in Doppelhofer & Weeks's (2009) research, (hereafter DW), who focused on the use of model averaging when jointness of the covariates is considered. DW's application is identical to FLS, studying the determinants of economic growth. Appendix B of their paper provides the BMA PIPs which I try to replicate using the ensemble of BMA packages. The data used in DW comprises 88 countries and 67 candidate variables as a cross section for the period 1960 to 1996. The definition of all 67 variables used can

²⁰This dataset is taken from the larger dataset used by Sala-i-Martin (1997) for his study on robust determinants of growth. The exact FLS dataset is publicly available on the *Journal of Applied Econometrics* online data archive.

²¹Both **BAS** and **BMA**, however, are computationally much faster than the **BMS** package.

Table A.8: Performance of BMS, BAS, and BMA Packages for the Fernandez et al. (2001*b*) Study.

	Published Results	BMS Package	BAS Package	BMA Package
	PIP	PIP	PIP	PIP
	(i)	(ii)	(iii)	(iv)
Variable:				
GDP 1960	1.000	1.000	1.000	1.000
Fraction Confucian	0.995	0.990	1.000	1.000
Life Expectancy	0.946	0.926	0.998	1.000
Equipment Investment	0.942	0.925	0.997	0.972
Sub-Saharan Dummy	0.757	0.747	*1.000	*1.000
Fraction Muslim	0.656	0.634	0.487	0.519
Rule of Law	0.516	0.510	*0.976	*1.000
Years Open Economy	0.502	0.502	*0.048	*0.055
Degree of Capitalism	0.471	0.485	0.460	*1.000
Fraction Protestant	0.461	0.450	0.355	*0.945
Fraction GDP in Mining	0.441	0.450	*0.978	*0.966
Non-Equipment Investment	0.431	0.441	*0.746	*0.987
Latin American Dummy	0.190	0.200	*0.884	*0.637
Primary School Enrollment, 1960	0.184	0.192	0.818	dropped
Fraction Buddhist	0.167	0.190	0.222	*0.551
Black Market Premium	0.157	0.177	*0.673	*0.841
Fraction Catholic	0.111	0.124	0.083	dropped
Civil Liberties	0.100	0.112	*0.733	dropped
Fraction Hindu	0.097	0.111	*0.991	*0.737
Primary exports, 1970	0.071	0.093	0.039	dropped
Political Rights	0.069	0.092	0.184	*0.419
Exchange Rate Distortions	0.060	0.077	0.087	dropped
Age	0.058	0.075	0.154	0.210
War Dummy	0.052	0.074	0.108	dropped
Size Labor Force	0.047	0.052	*0.972	*0.567
Fraction Speaking Foreign Language	0.047	0.065	0.042	dropped
Fraction of State Population. Speaking English	0.047	0.067	*0.496	0.241
Ethnolinguistic Fractionalization	0.035	0.057	*0.933	dropped
Spanish Colony Dummy	0.034	0.047	*0.785	0.339
S.D. of Black-Market Premium	0.031	0.047	0.033	dropped
French Colony Dummy	0.031	0.042	*0.745	0.207
Absolute Latitude	0.024	0.039	0.045	dropped
Ratio Workers to Population	0.024	0.040	0.032	dropped
Higher Education Enrollment	0.024	0.039	*0.936	dropped
Population Growth	0.022	0.040	0.049	dropped
British Colony Dummy	0.022	0.040	*0.651	0.053
Outward Orientation	0.021	0.036	*0.554	dropped
Fraction Jewish	0.019	0.030	0.034	dropped
Revolutions and Coups	0.017	0.027	0.032	dropped
Public Education Share	0.016	0.028	0.196	dropped
Area	0.016	0.020	0.032	dropped
Computation Time:				
(Seconds)	—	140	46	10

Notes: This table presents the results of applying Bayesian model averaging to the data used in Fernandez et al. (2001*b*) using BMS, BAS, and BMA packages. The dependent variable is the growth rate from 1960-1996 across 72 countries. All models estimated contain an intercept. Column (i) shows the Bayesian posterior inclusion probabilities reported by Fernandez et al. (2001*b*, Table I). Columns (ii), (iii), and (iv) report the posterior inclusion probabilities (PIP) using BMS, BAS, and BMA packages, respectively.

be found in data appendix B in DW and the dataset is publicly available on the *Journal of Applied Econometrics* data archive. As before I have tried to preserve the setup in DW by setting the packages' options as identical as possible.

Table A.9: Performance of BMS, BAS, and BMA Packages for the Doppelhofer & Weeks (2009) Study.

	Published Results	BMS Package		BAS Package		BMA Package	
	PIP (i)	PIP (ii)	Post Mean & SD (iii)	PIP (iv)	Post Mean & SD (v)	PIP (vi)	Post Mean & SD (vii)
Variable:							
East Asian Dummy	0.828	0.811	0.0219 (0.0126)	0.999	0.0227 (0.0071)	1.00	0.0297 (0.0065)
Primary Schooling Enrollment	0.796	0.799	-0.00006 (0.00004)	1.00	-0.00009 (0.00002)	1.00	-0.00008 (0.00002)
Investment Price	0.774	0.780	0.0167 (0.0101)	0.071	0.0003 (0.0022)	1.00	0.0168 (0.0049)
Initial GDP 1960	0.685	0.701	-0.0065 (0.0045)	0.999	-0.0085 (0.0026)	1.00	-0.0090 (0.0025)
Fraction of Tropical Area	0.570	0.565	-0.0082 (0.0079)	0.236	-0.0211 (0.0033)	1.00	-0.0155 (0.0030)
Computation Time: (Seconds)			160		44		20

Notes: This table presents the results of applying Bayesian Model Averaging to the data used in Doppelhofer & Weeks (2009) using BMS, BAS, and BMA packages. The dependent variable is the growth rate from 1960-1996 across 88 countries. Standard errors are in parentheses. All models estimated contain an intercept. Column (i) shows the Bayesian posterior inclusion probabilities reported by Doppelhofer & Weeks (2009, Data Appendix B). Columns (ii), (iv), and (vi) report the posterior inclusion probabilities (PIP) using BMS, BAS, and BMA packages, respectively. Columns (iii), (v), and (vii) show the posterior estimates of model coefficients and their standard errors using the three packages. The bottom part of table shows the computation time in seconds.

Table A.9 displays my findings. Column (i) shows the published PIPs > 0.50 that DW report in Appendix B of their paper. The rest of table presents the PIPs, posterior means and standard deviations from each of the packages.²² The results indicate that the BMS package is the only one that successfully reproduces the reported PIPs and posterior mean/standard deviation. Both the BAS and BMA packages reasonably reproduce the posterior means/standard deviations but the computed PIPs are significantly different from the published PIPs. For instance, the probability that “Investment Price” belongs to the final model is roughly 77% according to DW but the BAS packages reports this probability at nearly 7% and the BMA package reports it at exactly 100%!

The estimated coefficients are reasonably close for all packages yet there remain some anomalies. The estimated posterior mean on Fraction of Tropical Area in the BMS package is about a third as small as that reported from the BAS package and half the size of the reported posterior mean from the BMA package. Moreover, both the BAS and BMA packages are suggestive that Fraction of Tropical Area is relevant from a pure *t*-ratio perspective (see Masanjala & Papageorgiou 2008). Beyond differences in several of the posterior means across the packages the standard deviations show noticeable differences; compare the results for Investment Price where the standard deviation from the BMS package is nearly double that from the BMA package and almost five times as large from the reported standard

²²DW do not report estimated posterior means or standard deviations.

deviation in the **BAS** package.

A.8 Conclusions

This paper has outlined the currently available BMA packages (**BMS**, **BAS**, and **BMA**) in the statistical computing environment R. My goal was to familiarize users with the different options that the current versions of the packages have to offer. I highlighted how each of the packages implements a BMA analysis as well as the options available to the user and the outputs that are returned. To further cement the operation of these packages and to determine how similar the packages are in practice, I presented a simple empirical example that first allowed all three packages to fully enumerate the model space. Beyond this I enhanced my empirical example to force all three packages to engage in search mechanisms throughout the model space. When the model space is relatively small, we see that all three packages are successful at matching the PIPs, posterior means, and posterior standard deviations. However, for the larger model space similarity of the PIPs broke down considerably.

To further buttress my investigation and comparison of these packages I also compared runtimes of generic calls to each package for a range of covariate and sample sizes. In most instances the **BAS** package was the fastest, especially for large problems (both in terms of the number of covariates and the number of observations). Additionally, I also sought to replicate two recent studies that deployed BMA to investigate the determinants of economic growth. Both of these studies used high level programming outside of R and as such represent the perfect opportunity to see how well these freely available packages compare to computer code specifically tailored to the problem at hand. My results were striking. The **BMS** package almost exactly reproduced the results from both studies while the **BAS** and **BMA** packages were not able to match the reports PIPS in either study but were reasonably accurate at constructing the posterior means and standard deviations of my second study (compared with the same estimates from the **BMS** package).

In sum, it appears that while the **BMS** package is invariably slower than its peers, its numerous options and flexibility suggest that it should make its way into the toolkit of applied researchers seeking to use BMA in their analysis. The results from the empirical examples from published studies suggest that while both the **BMS** and **BAS** packages offer a similar array of options, the **BMS** package is capable of replicating published studies deploying BMA at the cost of slightly longer run times. My apparent advocacy of the **BMS** package does not hinge on its ability to reproduce the results of published studies however, as this presumably just means that the original authors used an implementation similar to that of the **BMS** package which the other packages were unable to match. This in no way is an indicator of superiority. Lastly, the relative rigidity of the **BMA** package to that of both **BMS** and **BAS** suggests that its use in applied work should be carefully scrutinized.

Appendix B

Computer Codes

Listing B.1: Computer codes used in Appendix A

```
2 # Loading datasets and required libraries
3 #####

5 set.seed(2011)
6 library(MASS)
7 library(tikzDevice)
8 data(UScrime); attach(UScrime)
9 UScrime1 <- (cbind(log(UScrime[,c(16,1,3:15)]), So))
10 noise<- matrix(rnorm(35*nrow(UScrime)), ncol=35)
11 colnames(noise)<- paste('noise', 1:35, sep='')
12 UScrime.log <- cbind(UScrime1, noise)
13 X <- UScrime1[, -1]; Y <- UScrime1[, 1]
14 x <- UScrime.log[, -1]; y <- UScrime.log[, 1]

17 # Full enumeration of model space (Tables A.3 and A.4)
18 #####

20 # BMS Package:
21 library(BMS)
22 bms_enu <- bms(UScrime1, mcmc="enumerate", g="UIP", mprior="
    uniform", user.int=FALSE)
23 coef(bms_enu)
24 detach("package:BMS")

26 # BMA Package:
27 library(BMA)
28 bma_enu <- iBMA.bicreg(X, Y, thresProbne0 = 5, verbose = TRUE,
    maxNvar = 30)
29 summary(bma_enu)
30 detach("package:BMA")

32 # BAS Package:
33 library(BAS)
```

```

34 bas_enu<- bas.lm(y~., data=UScrime1, n.models=NULL, prior="ZS-
    null", modelprior=uniform(), initprobs="Uniform")
35 coef(bas_enu)
36 detach("package:BAS")

39 # Using model space search mechanisms (Tables A.5 and A.6)
40 #####

42 # BMS Package:
43 library(BMS)
44 bms.time <- system.time(bms_sam <- bms(UScrime.log, burn=2000,
    iter = 100000, mcmc="bd", g="UIP", mprior="uniform", nmodel
    =2000, user.int=FALSE))
45 cat("Elapsed Time=", bms.time[3], "Seconds")
46 # Figure 1
47 plot(bms_sam[1:200])
48 # Figure 2
49 bms_sam2 <- bms(UScrime.log, burn=2000, iter = 100000, mcmc="bd",
    g="UIP", mprior="fixed", nmodel=2000, user.int=FALSE)
50 plot(bms_sam2[1:200])
51 # Figure 3(a)
52 density.bma(bms_sam, reg = "GDP", addons="esEpl")
53 detach("package:BMS")

55 # BMA Package:
56 library(BMA)
57 bma.time <- system.time(bma_sam <- iBMA.bicreg(x, y, thresProbne0
    = 5, verbose = TRUE, maxNvar = 30, nIter = 100000))
58 cat("Elapsed Time=", bma.time[3], "Seconds")
59 #Figure 3(b)
60 bma_fig <- bicreg(x, y)
61 plot(bma_fig, include=5, include.intercept=FALSE)
62 detach("package:BMA")

64 # BAS Package:
65 library(BAS)
66 bas.time <- system.time(bas_sam <- bas.lm(y~., data=UScrime.log,
    n.models=2^15, prior="ZS-null", modelprior=uniform(),
    initprobs="Uniform", method="AMCMC", Burnin.iterations=2000,
    MCMC.iterations=100000))
67 cat("Elapsed Time=", bas.time[3], "Seconds")
68 # Figure 3(c)
69 plot(coefficients(bas_sam), subset=12, ask=TRUE)
70 # Figures 4(a), 4(b), and 4(c)

```

```

71 plot(bas_sam, add.smooth=TRUE, panel="add.smooth")
72 detach("package:BAS")

75 # Evaluating time performance of the packages (Table A.7)
76 #####

78 UScrime.time <- as.data.frame(matrix(rnorm(6050000), ncol=121))

80 # BMS Package:
81 library(BMS)
82 bms.mat <- matrix(0,4,4)
83 m <- 1; n <- 1
84 for(i in c(15,30,60,120)){
85   for(j in c(50,500,5000,50000)){
86     bms.run.time <- system.time(bms.run <- bms(UScrime.time[1:
      j,1:(i+1)], burn=2000, iter = 100000, mcmc="bd", g="UIP
      ", mprior="uniform", nmodel=2000, user.int=FALSE))
87     bms.mat[m,n] <- bms.run.time[3]
88     m <- m+1
89   }
90   m <- 1
91   n <- n+1
92 }
93 detach("package:BMS")

95 # BMA Package:
96 library(BMA)
97 bma.mat <- matrix(0,4,4)
98 m <- 1; n <- 1
99 x <- UScrime.time[, -1]
100 y <- as.data.frame(UScrime.time[, 1])
101 for(i in c(15,30,60,120)){
102   for(j in c(50,500,5000,50000)){
103     bma.run.time <- system.time(bma.run <- iBMA.bicreg(x[1:j
      ,1:i], y[1:j,], thresProbne0 = 5, verbose = TRUE,
      maxNvar = 30, nIter = 100000))
104     bma.mat[m,n] <- bma.run.time[3]
105     m <- m+1
106   }
107   m <- 1
108   n <- n+1
109 }
110 detach("package:BMA")

```

```

112 # BAS Package:
113 library(BAS)
114 bas.mat <- matrix(0,4,4)
115 m <- 1; n <- 1
116 for(i in c(15,30,60,120)){
117   for(j in c(50,500,5000,50000)){
118     bas.run.time <- system.time(bas.run <- bas.lm(V1~., data=
      UScrime.time[1:j,1:(i+1)], n.models=2^15, prior="ZS-
      null", modelprior=uniform(), initprobs="Uniform",
      method="AMCMC", Burnin.iterations=2000, MCMC.iterations
      =100000))
119     bas.mat[m,n] <- bas.run.time[3]
120     m <- m+1
121   }
122   n <- n+1
123 }
124 }
125 detach("package:BAS")

128 # Evaluating the replicability of three packages (Table A.8)
129 #####

131 mydata <- read.dta("fls21.dta", convert.dates=TRUE, convert.
  factors=TRUE, missing.type=TRUE, convert.underscore=TRUE, arn.
  missing.labels=TRUE)

133 bmadata <- mydata[, -2]
134 y <- bmadata[, 1] ; x <- bmadata[, -1]
135 attach(bmadata)

137 library(BMS)
138 bms.time <- system.time(fls.bms <- bms(bmadata, burn=10000, iter
  =500000, g="UIP", mprior="uniform", nmodel=3000, mcmc="bd",
  user.int=FALSE))
139 coef(fls.bms)
140 detach("package:BMS")

142 library(BAS)
143 bas.time <- system.time(fls.bas <- bas.lm(gr56092~., data=bmadata
  , n.models=2^15, prior="ZS-null", modelprior=uniform(),
  initprobs="Uniform", method="AMCMC", Burnin.iterations=10000,
  MCMC.iterations=500000))
144 coef(fls.bas)
145 detach("package:BAS")

```

```

147 library(BMA)
148 bma.time <- system.time(fls.bma<- iBMA.bicreg(x, y, thresProbne0
    = 5, verbose = TRUE, maxNvar = 30, nIter = 500000))
149 summary(fls.bma)
150 detach("package:BMA")

153 # Evaluating the replicability of three packages (Table A.9)
154 #####

156 # To produce Table A.9, replace FLS dataset in the former
157 # chunk with DW dataset available on the Journal of
158 # Applied Econometrics Data Archive and rerun the code

160 ##### << END >> #####

```
