Essays in Experimental and Development Economics

by

Andrew J. Healy

B.S. Applied Mathematics / B.A. Political Science (1998)
Yale University

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2005

Signature of Author ....
Department of Economics
15 July 2005

Certified by ....
Esther Duflo
Professor of Economics
Thesis Supervisor

Accepted by ....
Peter Temin
Elisha Gray II Professor of Economics
Chairman, Departmental Committee on Graduate Studies

Essays in Experimental and Development Economics

by

Andrew J. Healy

Submitted to the Department of Economics
on 15 July 2005, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis is a collection of three essays on economic development and experimental economics.

In Chapter 1, I present experimental evidence about how Thais treat information from domestic and foreign sources. Thai students answer a series of objective general-knowledge questions, both before and after observing answers given by American students and other Thai students. By looking at how subjects update their original answers after observing information, it is possible to estimate the weights that subjects assign to themselves, the American answers they see, and the Thai answers they see. Consistent with previous studies, I find that Thais exhibit a significant level of overconfidence in that they overweigh their initial answers. Despite their overconfidence, the relative weight that they give to answers given by Americans compared to answers given by other Thais is, in most instances, statistically indistinguishable from the optimal solution. Moreover, the experimental design allows me to distinguish between two possible explanations for this fact. Under one hypothesis, subjects overestimate the relative precision of American answers, but fail to recognize the value of independence, and the two errors cancel each other out. Under a second hypothesis, subjects recognize the relative accuracy of each group and appreciate the value of independence. The data rejects the first hypothesis and supports the second.

In Chapter 2, I report the results of an experiment that tests for the presence of an information endowment effect. Experimental evidence suggests that an individual who is endowed with a coffee mug or chocolate bar demands a much higher price to sell than an unendowed person is willing to pay to acquire the same good. This study shows that a similar phenomenon does not exist when the endowment consists of information rather than goods. The results suggest that the endowment effect operates primarily on preferences as opposed to judgment.

In Chapter 3, my co-author and I use econometric methods to create a data set that makes it possible to better identify what areas of Thailand are poor and unequal. We show the potential for our results to improve policies targeted at poor households.

Thesis Supervisor: Esther Duflo
Title: Professor of Economics

To Chinda,

my example in life.

Acknowledgments

I thank my advisors, Victor Chernozhukov, Esther Duflo, and Sendhil Mullainathan for their advice throughout this project. Special thanks go to Adis Israngkura, Chayun Tantisidivakarn, and Somchai Jitsuchon for helping me to arrange the experiment reported in Chapter 1. Thanks also to Ken Chetnakarnkul and Phasook Mengkred for dedicated and superb research assistance. For the research in Chapter 2, I thank James Konow for his help in running the experiment. The work in Chapter 3 would not have been possible without the help of Peter Lanjouw and Khuankaew Varakornkarn. K. Nuntaporn and K. Tuang at the Thailand Development Research Institute provided invaluable assistance at various points in this research. For helpful discussions relating to Chapter 1, I acknowledge and thank Dan Ariely, Ernst Fehr, Jerry Hausman, Dean Karlan, John List, Whitney Newey, Al Roth, Shawn Cole, Brian Healy, Peter Hinrichs, Byron Lutz, Daniel Paravisini, Petia Topolova, Nancy Qian, Chinda Wongngamnit and participants in the MIT development and econometrics lunches. I thank David Reiley for his help with the experiment in Chapter 2. Somchai Jitsuchon and I thank Nanak Kakwani and Chris Elbers for their help with Chapter 3. I thank the Russell Sage Foundation and the MIT Shultz Fund for providing financial support for this project and the National Science Foundation for a graduate fellowship.

Contents

Chapter 1

How Does an Advisor's Nationality Affect the Weight Given to Her Advice? Experimental Evidence from Thailand

Summary 1 *It is sometimes assumed that people in some developing countries accord excessive weight to outside information, particularly information from the West. However, this apparent excess weight may in fact be optimal if people recognize the value of the independence in Western information. This paper reports an experiment designed to investigate this issue. Thai students answer a series of objective general-knowledge questions, both before and after observing answers given by American students and other Thai students. By looking at how subjects update their original answers after observing information, it is possible to estimate the weights that subjects assign to themselves, the American answers they see, and the Thai answers they see. I develop and estimate an econometric model of information aggregation to see how subjects would update their answers if they used information optimally. Consistent with previous studies, I find that Thais exhibit a significant level of overconfidence in that they overweigh their initial answers. Despite their overconfidence, the relative weight that they give to answers given by Americans compared to answers given by other Thais is, in most instances, statistically indistinguishable*

from the optimal solution. Moreover, the experimental design allows me to distinguish between two possible explanations for this fact. Under one hypothesis, subjects overestimate the relative precision of American answers, but fail to recognize the value of independence, and the two errors cancel each other out. Under a second hypothesis, subjects recognize the relative accuracy of each group and appreciate the value of independence. The data rejects the first hypothesis and supports the second.

1.1 Introduction

How people in developing countries incorporate the opinions of others when making decisions is acknowledged as a crucial issue. Previous research has looked at how information sharing within social networks affects technology adoption (Foster and Rosenzweig (1995), Duflo, Kremer, and Robinson (2004), Munshi (2004)). Information sharing is also vital in promoting public health (Dearden, Pritchett, and Brown (2004), Hsieh and Lin (1997)) and in helping individuals make better decisions about how to save their income (Duflo and Saez (2002)). This research generally shows that having neighbors who have the right information helps people make good decisions. At the same time, it is more difficult for those individuals who have neighbors with the wrong information to make good decisions (Miguel and Kremer (2004), Munshi and Myaux (2002)). An outside group's independent perspective could help to break the cycle of bad information circulated within poorly informed groups (Menon and Pfeffer (2003)).

Casual observation suggests that Thais accord a large weight to outside information, in particular Western information. Many Thais believe that this presumed behavior comes from a simple irrational preference for all things Western. The underlying idea is that Western information receives more weight than it should. However, this apparent excess weight could actually be optimal if Thais appreciate the value of independence in Western information. Do people in developing countries appreciate the value of considering an independent perspective when making decisions? If Thais put high weight on Western information, for example, could this in part reflect an appreciation of the value of its independence from a Thai's viewpoint? Do Thais exhibit overconfidence in the sense that they do not listen enough to any source of information? This paper reports the results of an experiment designed to answer these questions.

In the first stage of the experiment, American and Thai introductory economics students answered general-knowledge questions in three broad categories: 1) questions about Bangkok or Thailand, 2) questions about Boston or the US, and 3) questions about both of them. Each question has a correct answer. For example, one question asked subjects about the average high daily temperature in January in Bangkok, the average high daily temperature in January in Boston, and the sum of those temperatures. Subjects answered questions in a variety of areas including weather, economic data, and social indicators. In the second stage of the experiment, a separate set of Thai students first answered the same set of questions as the first stage students. They then observed randomly selected Thai and American answers from the first stage to help them update their answers. For any given question, a subject observed information about the Bangkok part, the Boston part, the sum part, or all three. Figure 1 shows an example of what one group of subjects saw for the question about January temperature. These subjects saw that one Thai student thought the high January temperature in Boston was $23°C$, another Thai thought the answer was $20°C$, and that an American student thought the answer was $2°C$. By looking at how subjects use the answers they observe to update their initial answers, this experimental design provides estimates of the weights that subjects applied to observed American answers, to observed Thai answers, and to their own initial answers.

In addition, the empirical distribution of answers provides estimates of the weights that subjects would optimally use to minimize the mean-squared error of their final answers.[1] Both an econometric model and a non-parametric test are used to characterize optimal behavior. The optimal weights are a function of the mean-squared errors (MSE) of Americans and Thais and the shares of MSE that comes from group biases. If the Thai bias share is large, for example, Thais tend to make the same kinds of mistakes. When this is the case, the independent American perspective is especially valuable. A Thai subject, and the other Thais she observes, generally think it is fairly warm in Boston in January and they need an outside voice to correct this mistake. *Even when any one Thai answer is equally good as any one American answer,* the fact that the error has a group component means that an optimizing Thai should assign higher

[1] In later sections, a simple model of Bayesian updating is used to derive the optimal weights individuals should use when processing the information they observe. It it shown that this model's predictions are not sensitive to parametric assumptions. See Genest and Schervish (1985), West (1992), and West and Crosse (1992) for more general models of Bayesian updating.

weight to American answers than to other Thai answers. The American answers, which do not contain the bias that she shares with other Thais, are more valuable to a Thai subject. The questions about both Bangkok and Boston illustrate this idea. In terms of MSE, Americans and Thais are about equally good at answering these questions. Nevertheless, I show that the presence of group bias means that an optimizing Thai should put more than twice as much weight on an observed American answer than an observed Thai answer.

For the questions about both Boston and Bangkok (the sum questions) and those that refer to Boston only, the hypothesis that subjects apply the optimal weight ratio to weigh the observed American answers compared to the observed Thai answers cannot be rejected. Based on how good each group is at answering the questions and the extent of group bias, subjects should put about five times as much weight on American answers for the Boston questions and about two-and-a-half times as much weight on American answers for the questions about both Bangkok and Boston. They come very close to that behavior, and equality cannot be rejected. For questions that refer to Bangkok only, subjects actually underweigh American answers. When the questions do not refer to Bangkok, subjects perform well despite showing a considerable amount of overconfidence through excessive weight given to their initial answers.[2] A large psychological literature shows that individuals exhibit overconfidence in a variety of environments (Griffin and Tversky (1992), Keren (1987), Gigerenzer, Hoffrage, and Kleinbolting (1991), Camerer and Lovallo (1999)). The fact that, in the experiment reported in this paper, overconfidence coexists with otherwise rational behavior fits with Camerer and Lovallo's (1999) conclusions about their subjects.[3] I show that the Thais in the experiment could decrease their MSE only slightly by changing how they relatively weigh observed Thai compared to observed American information, but they could decrease it substantially by reducing their overconfidence.

Moreover, the experimental design makes it possible to show that, when the questions do not refer specifically to Bangkok, subjects treat the information they observe in a sophisticated way. To a Thai subject, American information has extra value due to its independence from

[2] A test reveals that anchoring is not the cause of my finding of severe overconfidence. This test involves a treatment group that did not provide initial answers before observing the randomly selected American and Thai answers.

[3] Discussing their results from an experiment in which individuals opened businesses in the lab, Camerer and Lovallo (1999) states "This is not to say that subjects behave irrationally - indeed, they forecast the number of competitors quite well, and pass most tests of expectational rationality. They are simply overconfident."

her perspective. This value comes from the fact that Thais tend to make the same kind of mistake and Americans do not share that group bias. It is possible to reject the hypothesis that subjects fail to recognize the fact that each group makes a different kind of mistake. In contrast, the hypothesis that subjects appreciate how good each group is at answering the questions and that each group makes its own kind of mistake can explain the data quite well. Despite their overconfidence, subjects show remarkable sophistication when deciding how to relatively weight the American information they see compared to the Thai information they see. The experimental results support the appropriateness of modeling economic agents who are irrational in one dimension, but rational in others.

Section 2 describes the experimental design. In Section 3, I provide a framework for analyzing information aggregation. Section 4 contains the summary statistics that give the optimal relative American-to-Thai weights that subjects should apply. In Section 5, I report the estimates of how subjects weigh their private signals and the answers that they observe, including a variety of tests for different behaviors. Using the nature of a subset of questions, Section 6 tests the hypothesis that subjects do not account for the fact that they tend to make the same mistakes as other Thais. Section 7 looks at how subjects could have improved their performance by changing their behavior in various ways. Section 8 concludes.

1.2 Experimental Design

Below is a description of the questionnaire that the students answered, the random selection of answers for the main treatment group, and two special groups included in the experiment.

1.2.1 Questions

The set of fifteen questions covered a range of topics. Thirteen of these questions had three parts. For example, one question asked about Thai and American political leaders. One part asked for the number of Thai prime ministers since 1960, another part asked for the number of American presidents since 1960, and the third part asked for the sum of those two numbers. One other question asked about the number of 25-29 year-olds with some university education. Another asks about January temperature. Appendix 1 contains the entire questionnaire, showing the format for all of the questions.

The presence of a group of questions makes it possible to estimate in a general way how much of the mistake that a group member makes comes from the common group effect. For example, Thais erroneously answer high as a group for the January Boston temperature. By looking across questions, the experiment can generally estimate what share of the mistakes that people make is attributable to the group and how much is individual variation.[4] If, across questions, it is consistently the case that the group mean is distant from the truth and that there is little dispersion around the group mean, then the share of total MSE that can be attributed to the group is large.

Previous research has shown that subjects show overconfidence when asked to answer general knowledge questions and that overconfidence is amplified when the questions are difficult (Griffin and Tversky (1992)). In my experiment, the questions needed to be difficult. It was important that subjects have considerable uncertainty about the answers to the question, so that the provided information would be of value in resolving that uncertainty. This question difficulty could be expected to lead to substantial overconfidence. The experimental design makes it possible both to control for overconfidence and to look directly at how subjects treat observed American information compared to observed Thai information.

1.2.2 Stage 1: Creating a pool of American and Thai answers

In Stage 1 of the experiment, 116 introductory economics students at MIT and 130 introductory economics students at Thammasat University's Rangsit campus answered the series of questions, giving estimates of the population distribution of private signals for each question. Students had 15 minutes to answer the survey. In both countries, the questionnaire was given at the end of introductory economics classes. In Thailand, the questionnaire and instructions were given in Thai.[5]

Subjects received monetary rewards for answering accurately. For the American students, the top three performers on the entire set of questions received \$50 and the top fifteen performers on the individual questions received \$10. Among the Thai students, the top five performers

[4] Section 4 describes the simple procedure for calculating within-group correlation.

[5] Five translations and re-translations, as well as two pilots, were used to verify accuracy.

on the overall questionnaire received 1000 baht (approximately $25) and the top twenty on the individual questions received 200 baht. The additional rewards for the Thai students reflected the larger sample size.

Subjects also answered an optional personal survey before completing the questionnaire. The American students were asked to indicate their country of citizenship and the country where they attended high school, since some MIT undergraduates are citizens of other countries. Excluding this group of students from the sample has a negligible effect on the distribution of answers. Therefore, random selection of answers was based on the entire set of MIT students.

1.2.3 Stage 2: Subjects observe American and Thai information

In Stage 2, 300 economics undergraduates at Thammasat's Ta Pra Chan campus and master's economics students at the National Institute for Development Administration (NIDA) first received instructions (both read aloud and given in a packet) and then answered the 15 questions. Subjects were informed that they would receive 100 baht for participating and 20 baht for each question that they answered within a range of being correct. To encourage effort on all questions, the range was wider for more difficult questions, and subjects were informed of this. For example, subjects had to answer very closely for the question about Bangkok temperature to receive payment and much less closely for the question about US GNP. The goal of the incentives was to approximately provide subjects with the objective of minimizing the MSE of their answers, while keeping the instructions as simple as possible. In Section 7, it is shown that an individual who seeks to maximize her payments would act in the same way as a mean-squared error minimizer.

In Stage 2, subjects first answered all of the questions using Microsoft Excel in computer labs at NIDA and Thammasat. They directly answered the Bangkok/Thailand and Boston/US questions and the sum was calculated from those answers. After all subjects answered the questions, they received a second set of instructions. Subjects were told that they would observe randomly selected answers from MIT and Thammasat-Rangsit students who answered the same set of questions. These answers were provided in a separate packet. In the packet, for each question, subjects saw (as in Figure 1) the heading "Answers from Thai students" followed by the Thai information, and then "Answers from American students" followed by the

American information. Payments were based on subjects' final answers.

For each question, a subject saw information about either the Bangkok/Thailand question, the Boston/US question, the sum question, or all three, each with $\frac{1}{4}$ probability. Subjects in the main group saw up to three American answers and up to Thai answers. For each question, a subject could observe any of twenty randomly selected sets of American and Thai answers.

1.2.4 Special groups: Anchoring and large samples

Tversky and Kahneman (1974) shows that individuals will tend to stick to a number that is given to them, even when that number is irrelevant to the question at hand, a phenomenon they called anchoring. In my experiment, subjects first answer the questions and then update their answers based on what they observe. Thus, anchoring is a serious concern in this experiment; a subject provides her own answer to which she can anchor and that number is relevant to the final answer that she gives.

Due to these concerns, an additional 42 students observed randomly selected answers and answered the questions, without first providing their private signals. This provides an opportunity to derive a test for anchoring. In the experimental data, a subject who anchors would fail to sufficiently change the answer that she gives after observing information, even though her true belief changes. If anchoring is present, the 42 subjects who do not provide their private signals will choose final answers closer to the answers they observe, since they cannot stick to their initial answer purely due to the fact that they express it.

In addition, 50 of the 300 students in the main group observed large samples for two questions.[6] Subjects saw either 0, 5, 10, or 20 American and Thai answers. The model of information aggregation presented in the next section applies to any amounts of American and Thai information that subjects see. Subjects may treat large amounts of information differently, though, either due to difficulties with processing all the information or a lower bound to the weight a subject will give to herself. The presence of large samples makes it possible to see how subjects behave when they see large amounts of information.

[6] These are questions 4 (high daily January temperature) and 8 (number of prime ministers/presidents).

1.3 Model of Information

1.3.1 The importance of the group effect

Described here is a general model of information aggregation that shows how subjects will weigh the information they see, under the assumption that they are mean-squared error (MSE) minimizers. Each question type (Bangkok/Thailand, Boston/US, or sum) is considered separately.

For a question q, take individual i in group j to have a private signal x_{ijq} about the correct answer for the question. The mean-squared error Δ^2_{jq} for a group j for question q is then

$$\Delta^2_{jq} = \frac{1}{N_j} \sum_{i=1}^{N_j} (x_{ijq} - Truth_q)^2 \ , \tag{1.1}$$

where N_j is the number of group j members in the sample and $Truth_q$ is the correct answer for question q. A group that is comparatively better at answering the question will have a lower MSE for the question. The population distributions obtained in Stage 1 will give estimates for Δ^2_{Aq} (the mean squared error for Americans) and Δ^2_{Tq} (the mean squared error for Thais) for each question q.

The group MSE can be broken down into the sample variance for the group (s^2_{jq}) and the squared group bias (α^2_{jq}), where μ_{jq} is the population mean for group j for question q.

$$\begin{aligned}
\Delta^2_{jq} &= \frac{1}{N_j} \sum_{i=1}^{N_j} \left[(x_{ijq} - \mu_{jq})^2 + (\mu_{jq} - Truth_q)^2 \right] \\
&= \frac{1}{N_j} \sum_{i=1}^{N_j} (x_{ijq} - \mu_{jq})^2 \ + \ (\mu_{jq} - Truth_q)^2 \\
&= s^2_{jq} + \alpha^2_{jq}
\end{aligned}$$

The presence of the group bias term accounts for the fact that individuals in group j may make the same kind of mistake in answering question q.

Define the share of a group's MSE that comes from the group effect by ρ_{jq}.

$$\rho_{jq} = \frac{\alpha^2_{jq}}{\Delta^2_{jq}} \tag{1.2}$$

For group j for question q, ρ_{jq} captures what share of total MSE comes from the common group bias. If individuals in group j make different kinds of mistakes from other individuals in their group, ρ_{jq} will be low, as MSE will primarily be caused by the sample variance. If people make the same kind of mistake, ρ_{jq} will be high, since group bias will cause most of the group's MSE.

I focus on three parameters of interest. The first is the American-to-Thai MSE ratio, $\frac{\Delta_{Aq}^2}{\Delta_{Tq}^2}$. When this is high, Americans are less accurate at answering the questions relative to Thais. The other two parameters of interest are the American and Thai group bias shares. Given the description of subject behavior in the following subsection, of primary interest are the averages of these parameters across questions. Where N_q is the number of questions,

$$\frac{\Delta_A^2}{\Delta_T^2} = \frac{1}{N_q}\sum_{q=1}^{N_q}\frac{\Delta_{Aq}^2}{\Delta_{Tq}^2},\ \rho_A = \frac{1}{N_q}\sum_{q=1}^{N_q}\rho_{Aq},\ \text{and}\ \ \rho_T = \frac{1}{N_q}\sum_{q=1}^{N_q}\rho_{Tq}. \tag{1.3}$$

If the group bias share for Americans is low, for example, then multiple American guesses would provide more information about the true value. On the other hand, when the American group bias share is high, then a large group of American private beliefs is only slightly more informative than a single American private belief, because Americans tend to make the same kind of mistake. When the American (Thai) group bias share is greater than zero, it is rational to put lower weight on any one piece of American (Thai) information when more American (Thai) answers are available. Importantly, it is rational for a Thai subject to put lower weight on an observed Thai answer even when she observes only one Thai answer. As a Thai individual, she shares the same group bias with the observed Thai.

1.3.2 A model of subject behavior

Described here is a model of how subjects treat the information they observe. A Thai subject is unlikely to treat an observed Thai answer the same as her private signal. When problems are difficult, individuals have shown overconfidence in a wide variety of environments (Griffin and Tversky (1992), Keren (1987), Camerer and Lovallo (1999)). To account for this, a subject's own perceived MSE is modeled as a fraction c of another Thai's. Where Δ_{sq}^2 is a subject's perceived MSE for question q,

$$\Delta_{sq}^2 = c\Delta_{Tq}^2\ , \tag{1.4}$$

and overconfidence implies $c < 1$. This way of measuring overconfidence corresponds to much of the psychological literature. Overconfidence is often measured by looking at how small a subject's confidence interval is for a quantity given to what it should be (Cesarini, Sandewall, and Johanneson (2003)). The formulation of overconfidence here implies that a subject perceives her confidence interval to be c times the width of another Thai's, for any given confidence level.

It is assumed that a subject also perceives her bias to be the same fraction of another Thai's bias. Her perceived correlation with another Thai is then the same fraction c of the correlation between two other Thais.

$$\alpha_{sq} = c\alpha_{Tq} \Rightarrow E((x_{sq} - Truth_q)(x_{Tq} - Truth_q)) = c\rho_T\Delta^2_{Tq}. \tag{1.5}$$

A subject is assumed to minimize the expected MSE of her final answer. She weighs information according to her perceptions of how relatively good each group is at answering the type of question and how much of each group's MSE is captured by the group effect. Her behavior is a function of her *perceptions* of

$$\frac{\Delta^2_T}{\Delta^2_A}, \ \rho_T, \text{ and } \rho_A \ .$$

The optimal weights a subject would use are a function of the *actual* values of the parameters.

For the case where subjects see n_A American answers (x_{A1} through x_{An_A}) and n_T Thai answers (x_{T1} through x_{Tn_T}), a subject's objective function is:

$$E(MSE) = E\left(\widetilde{\beta}_T(x_{T1} + ... + x_{Tn_T}) + \widetilde{\beta}_A(x_{A1} + ... + x_{An_A}) + \widetilde{\beta}_s x_s - Truth\right)^2$$

where

$$\begin{aligned} \widetilde{\beta}_s &= \text{weight for own information} \\ \widetilde{\beta}_T &= \text{weight for any one piece of Thai information} \\ \widetilde{\beta}_A &= \text{weight for any one piece of American information} \end{aligned}$$

Assuming that $n_T\widetilde{\beta}_T + n_A\widetilde{\beta}_A + \widetilde{\beta}_s = 1$ so that the sum of the weights put on all pieces of information is one, this equation becomes

$$E(MSE) = E\left(\beta_T \sum_{k=1}^{n_T}(x_{Tk} - Truth) + \beta_A \sum_{k=1}^{n_A}(x_{Ak} - Truth) + \beta_s(x_s - Truth)\right)^2$$

Also assume that subjects perceive that the group biases are uncorrelated: $E((x_{Tk_T} - Truth)(x_{Ak_A} - Truth)) = 0$. The data does not reject the hypothesis that the actual group biases are uncorrelated.

$$\begin{aligned} E(MSE) &= n_T\beta_T^2\Delta_T^2 + n_T(n_T-1)\beta_T^2\rho_T\Delta_T^2 + n_A\beta_A^2\Delta_A^2 + n_A(n_A-1)\beta_A^2\rho_A\Delta_A^2 \\ &\quad + 2cn_T\beta_T\beta_s\rho_T\Delta_T^2 + c\beta_T^2\Delta_T^2 \end{aligned}$$

The first two terms in the above expression describe the expected MSE of all the observed Thai answers and the error that comes from the shared group bias among the n_T Thais (a total of $n_T(n_T-1)$ interactions). The next two terms capture the analogous errors for the observed American answers. The next-to-last term describes the shared group bias between the subject herself and the other Thais she observes. The final term describes the subject's own perceived MSE.

Taking the derivatives with respect to the weights gives the weights that subjects actually use and the optimal weights they should use to minimize MSE. The weights that subjects actually use are determined by subjects' perceptions of the parameter values. The optimal weights come from substitution of the actual parameter values into the same expressions, taking c as given. That is, the optimal weights will define optimal behavior given any level of overconfidence. The unconstrained optimal weights are found by using $c = 1$ (no overconfidence).

The actual weights are described by the following ratios.

$$\text{For } n_T \geq 1\text{: } \frac{\widetilde{\beta}_s}{\widetilde{\beta}_T} = \frac{1}{c} + \frac{1-c}{c}\left(\frac{(\rho_T)_{perceived}}{1-(\rho_T)_{perceived}}\right)n_T \tag{1.6}$$

$$\text{For } n_T = 0,\ n_A \geq 1\text{: } \frac{\widetilde{\beta}_s}{\widetilde{\beta}_A} = \left(\frac{\Delta_A^2}{\Delta_T^2}\right)_{perceived}\left(\frac{1-(\rho_A)_{perceived}}{c} + \frac{(\rho_A)_{perceived}}{c}n_A\right) \tag{1.7}$$

$$\text{For } n_T \geq 1,\ n_A \geq 1\text{: } \frac{\widetilde{\beta}_A}{\widetilde{\beta}_T} = \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{perceived}\left(\frac{1+(n_T-1)(\rho_T)_{perceived} - c(\rho_T)^2_{perceived}n_T}{(1+(n_A-1)(\rho_A)_{perceived})(1-(\rho_T)_{perceived})}\right) \tag{1.8}$$

The optimal weights are described by the same equations, with the actual parameter values substituted for the perceived values.

$$\text{For } n_T \geq 1\text{: } \frac{\widetilde{\beta}_s}{\widetilde{\beta}_T} = \frac{1}{c} + \frac{1-c}{c}\left(\frac{(\rho_T)_{actual}}{1-(\rho_T)_{actual}}\right) n_T \tag{1.9}$$

$$\text{For } n_T = 0,\ n_A \geq 1\text{: } \frac{\widetilde{\beta}_s}{\widetilde{\beta}_A} = \left(\frac{\Delta_A^2}{\Delta_T^2}\right)_{actual}\left(\frac{1-(\rho_A)_{actual}}{c} + \frac{(\rho_A)_{actual}}{c} n_A\right) \tag{1.10}$$

$$\text{For } n_T \geq 1,\ n_A \geq 1\text{: } \frac{\widetilde{\beta}_A}{\widetilde{\beta}_T} = \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{actual}\left(\frac{1+(n_T-1)(\rho_T)_{actual} - c(\rho_T)^2_{actual} n_T}{(1+(n_A-1)(\rho_A)_{actual})(1-(\rho_T)_{actual})}\right) \tag{1.11}$$

Equation (1.6) is the ratio of the self-weight to the weight for other Thais and thus the accuracy of Thais does not matter. But group bias does matter. When ρ_T is high and subjects are overconfident, the weight that subjects put on other Thais becomes small. When the group effect accounts for more of total MSE, an overconfident subject puts more trust in her reading of the joint Thai information than in another Thai's report. If $c = 1$, she puts equal weight on herself and another Thai.

Equation (1.7) is the ratio of the self-weight to the weight subjects put on observed Americans when $n_T = 0$. Subjects put more weight on Americans when Americans are perceived to be more accurate. This corresponds to subjects perceiving $\frac{\Delta_A^2}{\Delta_T^2}$ to be low. When c is low, subjects put less weight on American answers, since subjects perceive themselves to be better at answering the question. When ρ_A is high, subjects treat each additional American answer after the first as providing little added value. Dividing each side by n_A gives the optimal ratio of own-weight to total weight given to all American information.

$$\frac{\widetilde{\beta}_s}{n_A\widetilde{\beta}_A} = \left(\frac{\Delta_A^2}{\Delta_T^2}\right)_{perceived}\left(\frac{1-(\rho_A)_{perceived}}{cn_A} + \frac{(\rho_A)_{perceived}}{c}\right) \tag{1.12}$$

This shows more clearly that when the perceived ρ_A is high, subjects put less weight on each American answer when they observe more of them.

Equation (1.8) gives the weight ratio that subjects assign to an American answer relative to an observed Thai answer. Not surprisingly subjects put higher weight on American information when Δ_A^2 is low relative to Δ_T^2. Also when ρ_A is low and ρ_T is high, subjects put higher relative weight on American information.

The overconfidence parameter enters the expression in a second-order way through the $c\rho_T^2 n_T$ term. For reasonable values of ρ_T, overconfidence has a small, but noticeable, effect on the relative weight ratio that subjects use for American versus Thai information. When overconfidence is high (c is lower), subjects put more relative weight on Americans because a subject overweighs herself relative to a Thai both through (1.4) and (1.5), but only through (1.4) relative to an American. In other words, an overconfident subject trusts her reading of the common Thai information for a given question. Another way to think of this is that overconfident subjects are already putting high weight on Thai information through the weight they give to themselves.

One important detail to notice is that ρ_A only causes subjects to put less weight on individual American answers when n_A is greater than one, but that ρ_T causes subjects to put less weight on observed Thais even when only one Thai is observed. When one Thai is observed, there are two Thai answers to consider: a subject's own answer and the one she observes. As a result, the group effect enters (1.8) when $n_T = 1$, but not when $n_A = 1$.

The experimental data on how subjects update their answers provide estimates of the actual weights that subjects use. While the experiment does not directly observe subjects' perceptions of $\frac{\Delta_T^2}{\Delta_A^2}$, ρ_T, and ρ_A, Sections 5 and 6 will describe tests that make it possible to reject several possibilities for subject perceptions.

1.3.3 Non-structural optimal weight estimation

This section presents a series of non-structural tests for optimal behavior that complement the decision-making model's set of optimal weights. Checking how closely the non-structural estimates of the optimal weights match the model's estimates provides a test of the model's appropriateness.

Given that subjects minimize mean-squared error, the optimal weights that subjects should apply can be estimated from a simple regression. To run this regression, the data is first standardized to put each question on the same scale. Where the subscript i denotes the individual, j denotes the group (either A or T), and q denotes the question, any variable x_{ijq} is standardized in the following way[7]

$$z_{ijq} = \frac{x_{ijq} - \overline{x}_{Tq}}{\frac{1}{2}(s_{Tq} + s_{Aq})} \; .$$

In the above equation, s_{Tq} is the standard deviation of the Thai answers, s_{Aq} is the standard deviation of the American answers, and $\overline{x}_{Tq}$ is the mean of the Thai answers, for question q.

Where

$$\begin{aligned}
Truth_q &= \text{(standardized) correct answer for question } q, \\
\overline{z}_{iAq} &= \text{(standardized) average observed American answer for question } q \text{ seen by subject } i, \\
\overline{z}_{iTq} &= \text{(standardized) average observed Thai answer for question } q \text{ seen by subject } i, \\
z_{iq} &= \text{(standardized) initial answer for the subject who is choosing,}
\end{aligned}$$

the regression equation is

$$Truth_q = \alpha_A \overline{z}_{iAq} + \alpha_T \overline{z}_{iTq} + \alpha_s z_{iq} + \varepsilon_{iq} \tag{1.13}$$

Without parametric assumptions, this regression estimates the weights that a subject should use to minimize her mean-squared prediction error. This regression can be expanded to include additional terms, such as terms that account for the numbers of observed American and Thai answers, n_{iA} and n_{iT}). The results that come from estimating (1.13) will be discussed in Section 5, along with the estimation of actual subject behavior. To estimate actual behavior, I simply replace $Truth_q$ in (1.13) with the subject's final answer. This gives the weights that subjects actually assign to their initial answers, the American answers they observe, and the Thai answers they observe.

[7] Another possibility is to use the standard deviation of the Thai answers to standardize. For questions where Thais have a small standard deviation, this will cause the American mean to be far from zero, if the American and Thai means are different. These points can then have excessive influence on the regression. The results are not sensitive to alternative ways of standardizing the data.

1.4 Summary Statistics

The data on subjects' private signals shows that group bias is an important consideration. As an example, figures 2A, 2B, and 2C show kernel density estimates for the Thai and American answers for the questions about high daily January temperature.[8] The figures show that Thais have a mean answer of $20°C$ for the Boston January temperature (truth=$2°C$), and Americans have a mean of $20°C$ for the Bangkok temperature (truth=$32°C$). Across questions, knowing where the Thai mean is relative to the truth is uninformative about where the American mean is relative to the truth.

Based on the set of American and Thai answers for all of the questions, the Thai-to-American mean-squared error ratio can be estimated. If, as is estimated for the questions about Thailand, $\frac{\Delta_T^2}{\Delta_A^2}$ is 0.517, the expected squared distance between a randomly selected Thai answer and the truth is 0.517 times the expected squared distance between a randomly selected American answer and the truth.

Consider the questions of any type (Bangkok/Thailand, Boston/US, or sum). The ratio of the mean squared error for Thais relative to Americans can be calculated for each of the thirteen questions. The average across questions provides the MLE for the MSE ratio for the two groups for a given type of question:

$$\frac{\widehat{\Delta_T^2}}{\Delta_A^2} = \frac{1}{13}\sum_{q=1}^{13}\frac{\widehat{\Delta_{Tq}^2}}{\widehat{\Delta_{Aq}^2}}\ . \tag{1.14}$$

Table 1 summarizes the relative Thai-to-American accuracy for each of the three question types. The estimates in Table 1 come from the 116 American answers and the 430 Thais who provided their private signals.[9]

Thais have about one-half the MSE of Americans for the Thai questions. Americans are about three times more accurate for the questions about the US, and $\frac{\widehat{\Delta_T^2}}{\Delta_A^2}$ is about 1.3 for the

[8] These are Epanechnikov kernel estimates with optimally chosen bandwidth.

[9] Thai subjects were informed that the answers they observed came from MIT students and Thammasat-Rangsit students. So if subjects had different perceptions about Thammasat-Rangsit than the universe of all subjects, it would be appropriate to use only the 130 Thai students from Stage 1 to calculate variances and correlations. The ratios are nearly the same if just the Stage 1 students are used to estimate the variances and correlations.

sum questions. These ratios establish the relative weights that subjects would optimally apply if they ignored group bias. A Thai using $\frac{\Delta_T^2}{\Delta_A^2}$ to relatively weigh observed American answers versus observed Thai answers would ignore the fact that she tends to make the same mistakes as other Thais.

The experimental design also enables me to estimate the share of group bias in total MSE for each question type (Bangkok/Thailand, Boston/US, or sum). For each question, the realization of the group effect, α_{jq}, is used to estimate the bias for group j for question q.

$$\widehat{\alpha}_{jq} = \frac{1}{N_T}\left(\sum_{i=1}^{N_T} x_{iTq}\right) - Truth_q = \overline{x}_{Tq} - Truth_q \tag{1.15}$$

The share of the group effect (α_{jq}) in total mean squared error (Δ_{jq}^2) is

$$\phi_{jq}^2 = \frac{\alpha_{jq}^2}{\Delta_{jq}^2}. \tag{1.16}$$

The estimate of the group bias share for each question type is the mean of ϕ_{jq}^2 across questions:

$$\widehat{\rho}_j = \widehat{\phi_j^2} = \frac{1}{13}\Sigma_{q=1}^{13}\widehat{\phi}_{jq}^2 = \frac{1}{13}\Sigma_{q=1}^{13}\frac{\widehat{\alpha_{jq}^2}}{\widehat{\Delta_{jq}^2}}\ . \tag{1.17}$$

This captures the share of the group's total MSE for which the group effect is responsible.

Theorem 1: *Equation (1.17) is the MLE for $\widehat{\rho}_j$ under the model.*

Appendix 2 contains the proof for Theorem 1. Table 2 displays the group bias shares that (1.17) gives for each question type.

Importantly, the bias shares are larger for the question types about which each group has less knowledge. The share of total Thai mean squared error that comes from the group effect is smallest for the Bangkok/Thailand questions and largest for the Boston/US questions. In contrast, the share of total American mean squared error that comes from the group effect is smallest for the Boston/US questions and largest for the Bangkok/Thailand questions. For example, in the sample question in the introduction, Thais make small errors about the average high daily January Bangkok temperature and the group effect is a small share of that error. On the other hand, Thais make much larger errors when guessing the January Boston temperature, and a larger share of their mistakes comes from the fact that the group mean for Thais is $20°C$.

As will be shown in Section 6, this has important implications for the optimal weights for the case where subjects see information for the sum question only. If subjects correctly account for group bias when deciding how to treat the information they observe, this will have the same implications for the weights they actually use.

In summary, the answer distributions show significant group biases for both Americans and Thais. American answers thus have extra value to a Thai due to the independence of observed American answers compared to observed Thai answers from her perspective. An optimizing Thai subject needs to account for this when deciding how to weigh the American answers they see compared to the Thai answers they see.

1.5 Empirical Strategy

1.5.1 Estimating subject behavior

A simple regression provides the weights that subjects give to the information they observe and to their own private signals. This involves regressing subjects' final answers on their initial answers, the American answers they see, and the Thai answers they see. If subjects increase their answers more in response to high observed American answers than to high observed Thai answers, this will be reflected in a higher relative weight for the American answers.

The following regression estimates the average weights that a subject puts on American answers (β_A), other Thai answers (β_T), and her own private signal (β_s), modeling an individual's final choice as a linear combination of the average American answer she sees, the average Thai answer that she sees, and her own initial answer. Where

$$\begin{aligned}
y_{iq} &= \text{(standardized) choice of individual } i, \\
\overline{z}_{iAq} &= \text{(standardized) average observed American answer for question } q \text{ seen by subject } i, \\
\overline{z}_{iTq} &= \text{(standardized) average observed Thai answer for question } q \text{ seen by subject } i, \\
z_{iq} &= \text{(standardized) initial answer for the subject who is choosing,}
\end{aligned}$$

the regression equation is

$$y_{iq} = \beta_A \overline{z}_{iAq} + \beta_T \overline{z}_{iTq} + \beta_s z_{iq} + \varepsilon_{iq} \tag{1.18}$$

Equation (1.18) can be estimated with or without a constant and with or without assuming that the weights sum to one. Each of these corresponds to what would be expected of a rational agent. The first assumption states that if a subject gives an initial answer of zero and observes an American and Thai average of zero, she will give a final answer of zero. The second assumption states that a subject is assumed to put total weight of one on all information available to her. In all cases, these restrictions cannot be rejected. I focus on the unrestricted results because they have higher standard errors than the restricted results, which leads to more conservative hypothesis tests.

In the next subsection, I describe the weight estimates that come from (1.18). When subjects observe answers for the sum question only, it is important to note that subjects update their answers for both the Bangkok/Thailand and Boston/US questions. This means that there is more uncertainty about the answers that a subject sees; she does not see direct information about the answers she changes. As a result, it is not correct to compare estimates of β_s from (1.18) to optimal estimates based on the assumption that subjects update specifically for the sum question. Optimizing behavior in terms of choosing $\frac{\beta_A}{\beta_T}$ (for the sum question), though, is the same whether updating occurs directly for the sum question or the observed answers to the sum question are used to update for the Bangkok/Thailand and Boston/US questions. In other words, when subjects observe answers for the sum question only, regression (1.18) gives a shortcut for estimating $\frac{\beta_A}{\beta_T}$ for the sum question. Subject behavior on the Bangkok/Thailand and Boston/US questions when subjects observe answers for the sum question only will be considered in detail in Section 6.4.

Note that optimal behavior involves choosing different weights for different amounts, n_{iA} and n_{iT}, of observed information. When a subject observes more American answers, for example, a higher weight should be attributed to the American average since it contains more information. A high ρ_A implies that there is less new information in each additional American answer. When the group bias share is higher, a subject should put less additional weight on each American answer when she sees more of them.

While not necessary for subjects to apply the correct weights on average, if subjects have some knowledge of the group MSEs for individual questions, they will apply higher weight to American information when it is relatively better. To see if subjects behave this way, the regression can be expanded to include a term that captures group accuracy across questions:

$$Acc_q = \frac{\widehat{\Delta}^2_{Tq}}{\widehat{\Delta}^2_{Aq} + \widehat{\Delta}^2_{Tq}} \ . \tag{1.19}$$

When Acc_q is high, Americans are better relative to Thais for the question q.

Subjects may also want to account for the spread in the answers they observe. For example, when subjects see two answers that are near each other and one answer that is extreme, they may ignore the extreme answer. To look at this, the regression can be expanded to include a term that captures the spread in the observed answers. The spread is measured simply as the standard deviation in the observed information. Where x_{ijq} is the ith observation from group j for question q, this measure is

$$SD_{jq} = \frac{1}{N_j - 1} \sum_{i=1}^{N_j} \left(z_{ijq} - \frac{1}{N_j} \sum_{i=1}^{N_j} z_{ijq} \right)^2 \ . \tag{1.20}$$

Results when the SD_{Aq} and SD_{Tq} terms are included are available from the author upon request. Inclusion of these terms has no effect on the results to be discussed in the following sections.

The full regression allows each weight to vary with each term, so that the weights are modeled in the following way:

$$\beta_s = \beta_{s,1} + \beta_{s,2} n_{iA} + \beta_{s,3} n_{iT} + \beta_{s,4} Acc_q \tag{1.21}$$

$$\beta_A = \beta_{A,1} + \beta_{A,2} n_{iA} + \beta_{A,3} n_{iT} + \beta_{A,4} Acc_q \tag{1.22}$$

$$\beta_T = \beta_{T,1} + \beta_{T,2} n_{iA} + \beta_{T,3} n_{iT} + \beta_{T,4} Acc_q \tag{1.23}$$

The regression results indicate that subjects sometimes go beyond simply weighting the averages that they see. For all types of questions, they apply higher relative weights to American information for those questions on which Americans answer relatively well.

1.5.2 Estimated weights used by subjects

The equation (1.18) regressions give the estimated average weights that subjects put on Thai answers, American answers, and their own initial answers for each question type. Tables 3A, 3B, and 3C report the results. The cases where individuals see answers for all three types of questions are included in the regressions for the Thai and US questions. For the regressions for the sum question, these observations are left out. The assumption is that subjects reason based on the direct information that they see about the Thai and US questions.[10]

Table 3A shows the weights that subjects use when they observe both American and Thai information for the questions about Bangkok and Thailand. Consider the first column. Subjects put a weight of 0.654 on their private beliefs, 0.242 on the observed Thai average, and 0.058 on the observed American average. The model estimates that subjects assign 4.2 times more weight to the observed Thai answers than to American answers.

Table 3B shows the weights that subjects use when they observe both American and Thai information for the questions about Boston and the US. Again, consider the first column. The model estimates that subjects assign 5.1 times more weight to American answers than to other Thai answers, giving estimates of 0.461 for the own-weight, 0.092 for the Thai weight, and 0.466 for the American weight.

Table 3C shows the weights that subjects use when they observe both American and Thai information for the questions about the sum. The regression estimates that subjects assign 2.4 times more weight to American answers than to observed Thai answers, giving estimates of 0.733 for the own-weight, 0.069 for the Thai weight, and 0.167 for the American weight.

The subjects appear to account for different group accuracies across questions. For all three types of questions, as captured by the coefficients on Acc_q, subjects put significantly more

[10] One possibility is that there are spillovers across question types when subjects observe all answers for all three types of questions. For example, a subject who sees an American who answers well for the Thailand question may put more weight on that American for the US question. Testing for these generally produces insignificant results, although for the sum question, there is a small and significant spillover effect. Details are available upon request.

weight on American information for those questions on which Americans perform relatively better.

It may seem surprising that subjects are able to appreciate the accuracy of Americans relative to Thais for different questions. That they do is less surprising when the individual questions are considered. Thai students do much better relative to Americans for the temperature in Bangkok than they do for female-labor force participation in Thailand. It is not unreasonable for a Thai student to think that it takes individual experience in Bangkok to know the weather there, but that there is some general knowledge, that Americans may have as well as Thais, involved in answering the labor-force question. Therefore, the Thai subjects may and apparently do put higher relative weight on Thai information for the question about Bangkok weather.

1.5.3 Special groups

Large samples

The predictions of a statistically optimizing model of behavior and actual behavior are likely to deviate from each other as subjects observe increasing amounts of information, even if the model applies for smaller amounts of observed answers. In the extreme case, as the number of answers that a subject observes goes to infinity, the weight that a subject puts on her own private signal should go to zero. In practice, this is not likely to be the case. One major reason is the bounds to subjects' abilities to process information.

To see how larger samples could affect the way in which subjects use information, a treatment group of 50 subjects saw samples of 0, 5, 10, or 20 American and Thai answers for two questions.[11] Table 6 describes the share of subjects who leave their initial answer unchanged. For each type of question, subjects are significantly more likely to leave their initial answer unchanged when they observe more information.

One explanation involves subjects throwing up their hands due to difficulties in processing the observed information. Another possibility is that a subject who observes many answers is more likely to observe two that bracket her own, which a subject interprets as confirming her

[11] These are questions 4 (high daily January temperature) and 8 (numbers of political leaders) in the questionnaire.

initial answer.

Anchoring

In this experiment, anchoring is potentially a major concern. Subjects may adhere more closely to their initial answers because they have to express them. To test for this, a treatment group of 42 students observed randomly selected answers from Americans and Thais without first providing their private signals. If anchoring is present, subjects in this group should do better at answering the questions than those in the main group. The students in the anchoring treatment group will avoid the losses associated with anchoring to their initially expressed answer.

To test for this, I consider the following specification:

$$(y_{iq} - Truth_q)^2 = \beta_1 Anchor_i + \beta_{2v} Version_{iv} + u_{iq} \ , \tag{1.24}$$

where $Anchor_i$ is a dummy that is one if the subject is in the anchoring control group and zero otherwise. $Version_{iv}$ is a dummy that is one if subject i observed version v of the information out of the 20 that were used. If subjects suffer losses from anchoring, β_1 should be negative, so that subjects who do not have the opportunity to adhere to their stated answer achieve lower distances from the correct answer. Table 7 shows that this is the case, but not significantly so.

It is also possible to look at anchoring directly by comparing how close subjects' final answers are to the information they observe. If subjects anchor to their private signals when they reveal them, the subjects who do not reveal their private signals will choose final answers closer to the answers they observe.

To test for this, I consider the following regressions:

$$\begin{aligned} (y_{iq} - \overline{z}_{iTq})^2 &= \beta'_1 Anchor_i + \beta'_{2v} Version_{iv} + u'_{iq} \\ (y_{iq} - \overline{z}_{iAq})^2 &= \beta''_1 Anchor_i + \beta''_{2v} Version_{iv} + u''_{iq} \end{aligned} \tag{1.25}$$

Anchoring would imply that $\beta'_1 < 0$ and $\beta''_1 < 0$. Table 8 shows that this generally appears to be the case, although never significantly so. The last column shows the average mean-squared distance (from the observed Thai or American average) for the main group and this puts the regression coefficients in perspective.

For example, subjects in the main group choose final answers that are 1.835 standard deviations from the observed Thai average for the US question. The regression estimates that the students in the anchoring group who observe both kinds of answers choose final answers that are 0.051 standard deviations closer to the observed Thai average than the main group. In general, the estimated coefficients follow this pattern. The data shows that anchoring has a small and generally insignificant effect on subject behavior.

1.6 Tests for optimal behavior

1.6.1 Estimated optimal weights

Without any parametric assumptions, it is possible to estimate the optimal weights a subject should use by simply regressing the correct answer on a subject's initial answer, the average of the American answers she observes, and the average of the Thai answers she observes. The regression equation (1.13) in Section 3.3 expressed this idea. As was the case with the estimated actual weights from (1.18), these regressions can also be expanded to include terms with n_{iA}, n_{iT}, and Acc_q.Tables 4A, 4B, and 4C show the results.

The first column of each table captures the primary results. For the Bangkok/Thailand questions, an optimally-behaving subject would put a weight of 0.295 on her private belief, a weight of 0.458 on the observed Thai average, and a weight of 0.247 on the observed American average. For the Boston/US questions, the corresponding weights are 0.063, 0.115, and 0.822. For the sum questions, the regression estimates that a subject should choose 0.180 for the self weight, 0.250 for the weight given to observed Thai answers, and 0.570 for the weight given to observed American answers. Notice that these regression estimates of optimal behavior do not account for overconfidence. These are the weights that a subject should use assuming overconfidence is not present.

The optimal weights subjects should use can also be estimated using the econometric model described in Section 3. These estimates come from substituting the estimates of the American-to-Thai MSE ratio, the American group bias share, and the Thai group bias share into equations (1.9) and (1.11). The estimates of these parameters were reported in Tables 1 and 2.

Tables 5A, 5B, and 5C display the main results about the actual behavior and the two sets of optimal estimates to three comparable tables. Table 5C displays the econometric model's estimates of the optimal weights subjects should apply. For the Bangkok/Thailand questions, a subject should apply a weight of 0.258 to her initial answer, 0.459 to the Thai average she sees, and 0.283 to the American average she sees. For the Boston/US questions, the corresponding weights are 0.100, 0.174, and 0.726. For the sum questions, the model estimates that a subject would optimally choose 0.155 for the self-weight, 0.333 for the weight given to observed Thai answers, and 0.512 for the weight given to observed American answers. Confidence intervals for these estimates are obtained by bootstrapping.

For all three types of questions, equality between the regression's estimates and the econometric model's estimates of optimal behavior cannot be rejected. The close correspondence between the optimal weight estimates obtained with the regression and those that come from the econometric model provides additional evidence supporting the importance of accounting for group bias in determining optimal behavior. If, in the model, the group bias shares are assumed to be zero, the hypothesis of equality between the two sets of estimates of optimal behavior is rejected for each of the three types of questions.

1.6.2 Construction of confidence intervals

Confidence intervals can be constructed for the actual, naive, and optimal weight ratios. The actual weight ratio is estimated from subject behavior, as described in Section 5. The naive weight ratio expresses the relative weight a subject would use if she understood each group's accuracy and ignored group bias. The optimal weight ratio expresses the relative weight that a subject would use if she correctly perceived each group's accuracy and accounted for group bias.

Weight ratios:

$$\text{Estimated} = \frac{\beta_A}{\beta_T}$$

$$\text{Naive} = \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{actual}$$

$$\text{Optimal} = \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{actual} \left(\frac{1 + (n_T - 1)\,(\rho_T)_{actual} - c\,(\rho_T)_{actual}^2\, n_T}{(1 + (n_A - 1)\,(\rho_A)_{actual})(1 - (\rho_T)_{actual})}\right)$$

The confidence intervals for the optimal weight ratios are calculated for a variety of different values of the overconfidence parameter c. The optimal weight ratio increases as subjects have more overconfidence. More overconfident subjects believe they have a lower error variance relative to both Thais and Americans, but they also believe more in their reading of the shared Thai information about a question. This means that overconfidence has a larger negative effect on the weight subjects put on observed Thai information than on observed American information.

Even for very large overconfidence (small c), though, the direct effect of group bias on the optimal weight ratio is larger than the effect of overconfidence. Consider the Boston/US questions. With no overconfidence ($c = 1$), the group bias share causes the optimal weight ratio to be 4.423, compared to the naive weight ratio of 3.086. Increasing overconfidence by lowering c to 0.25 only causes the optimal weight ratio to rise to 5.125.

The data can also be used to estimate the overconfidence parameter with some additional assumptions. Specifically, for each of the question types, if it is assumed that $\rho_A = \rho_T$ and c is allowed to be a function of $n_{iA} + n_{iT}$, the overconfidence parameters $c(n_{iA} + n_{iT})$ can be estimated. Table 9 reports these results. The standard error bounds for the overconfidence estimates are large. Estimated overconfidence generally increases when more information is observed, but it is present for any amount of information. The estimated c values are used only for illustrative purposes in the calculation of the optimal American-to-Thai weight ratios.

1.6.3 A test for naive behavior

The experimental data can be used to test hypotheses about how subjects perceive the relative accuracy and group biases of Americans and Thais. Specifically, consider the hypothesis of naive behavior, N_0:

$$N_0 : \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{perceived} = \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{actual}, \ (\rho_T)_{perceived} = 0, \ (\rho_A)_{perceived} = 0 \ .$$

This hypothesis states that subjects correctly perceive the MSE of Thais relative to Americans, but ignore group bias. Under N_0, subjects understand how accurate each group is, but fail to value the independence in American information.

Under N_0, subjects will choose

$$\frac{\beta_A}{\beta_T} = \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{actual} . \tag{1.26}$$

A test of N_0 is a test of (1.26), equality between the actual weight ratio that subjects use and the naive weight ratio.

The second row of Table 10 displays the results of this test for all three types of questions. Confidence intervals for the actual weight ratios are calculated by bootstrapping the estimated regression coefficients using the estimated variance-covariance matrix. For the Bangkok/Thailand questions and the sum questions, we can reject this hypothesis at the 5% level ($p = 0$ and $p = 0.017$, respectively). For the Boston/US questions, we can reject it at the 10% level ($p = 0.058$). The hypothesis is rejected for the Bangkok questions due to subjects choosing too low a weight for American answers relative to Thai answers. It is rejected for the Boston/US and sum questions due to subjects relatively overweighing American answers. For all three types of questions, by looking at how subjects weigh the American answers they see compared to the Thai answers they see, it can be rejected that subjects correctly perceive how accurate Americans are relative to Thais while at the same time failing to recognize the presence of group bias.

Now consider the hypothesis of sophisticated behavior, S_0:

$$S_0 : \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{perceived} = \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{actual}, (\rho_T)_{perceived} = (\rho_T)_{actual}, (\rho_A)_{perceived} = (\rho_A)_{actual}$$

This hypothesis states that subjects correctly perceive the MSE of Thais relative to Americans, and also correctly account for group bias. Under S_0, subjects understand how accurate each group is and correctly value the independence in American information. Compared to a subject who behaves according to N_0, a subject who behaves according to S_0 will put more weight on American answers because she appreciates the value of an American's independent perspective.

Under S_0, subjects will choose the optimal weight ratio

$$\frac{\beta_A}{\beta_T} = \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{actual} \left(\frac{1 + (n_T - 1)(\rho_T)_{actual} - c(\rho_T)^2_{actual} n_T}{(1 + (n_A - 1)(\rho_A)_{actual})(1 - (\rho_T)_{actual})}\right) \tag{1.27}$$

A test of S_0 is a test of (1.27), equality between the relative weight ratio that subjects use and the optimal weight ratio.

Table 10 displays the results of this test for a variety of possible values of the overconfidence parameter. For the Bangkok/Thailand questions, we reject S_0. For all values of overconfidence, the p-value of the test is zero or nearly zero. Thais appear to put too little weight on American answers in this case.

On the other hand, for the Boston/US and sum questions, this hypothesis is not rejected for any level of overconfidence. For $c = 1$, the optimal weight ratio estimates are 4.423 and 1.843, compared to the actual weight ratio estimates of 5.065 and 2.420. The p-values for the two tests of equality are 0.610 and 0.367, respectively.

Given estimated overconfidence, the actual and optimal weight ratios match up remarkably closely in these two cases. For the Boston/US questions, the optimal weight ratio estimate is 5.043 and the actual weight ratio estimate is 5.065. The p-value for equality between the two is (not surprisingly) nearly one ($p = 0.963$). For the sum questions, the actual weight ratio estimate is 2.420, compared to the optimal estimate of 2.168, giving a p-value of 0.721 in the equality test. If individuals are overconfident, but still appreciate both how accurate Americans are relative to other Thais and the extent of group bias, they will use relative weight ratios that are very close to those actually used by the experimental subjects.

Still, the subject behavior described to this point for the Boston/US and sum questions can also be explained by a hypothesis of biased behavior, B_0:

$$B_0: \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{perceived} > \left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{actual}, \ (\rho_T)_{perceived} = 0, \ (\rho_A)_{perceived} = 0$$

Under B_0, subjects perceive Americans to be better than they actually are compared to Thais and they ignore group bias.

The experimental design makes it possible to distinguish between S_0 and B_0. The following subsection shows how the special nature of the sum question permits a test of the hypothesis that subjects ignore the group effect *for any value of* $\left(\frac{\Delta_T^2}{\Delta_A^2}\right)_{perceived}$, making it possible to distinguish between these explanations of subject behavior.

1.6.4 A test for sophisticated behavior

By looking at how subjects who observe answers only for the sum question update their answers for both the Bangkok/Thailand and Boston/US questions, it is possible to test the hypothesis that subjects ignore group bias when choosing their final answers. To derive this test, it is necessary to expand the earlier notation which applied when each question type was considered separately.

Define

$$\begin{aligned} \rho_{jk} &= \text{group bias share in total MSE for group } j \text{ for question type } k \\ \Delta^2_{jk} &= \text{mean-squared error for group } j \text{ for question type } k \ . \end{aligned}$$

For example, ρ_{Ta} is the group bias share for Thais for the Boston/US questions.

Consider the case when a subject uses observed answers for the sum question to update her answer for the Bangkok/Thailand question. A subject updates her answer for the Bangkok/Thailand question based on her initial answer for the Bangkok/Thailand question and the distance between the answers she observes for the sum question and her initial answer for the sum question. When a subject observes answers above her own for the sum question and uses them to update her answer, she is likely to revise upwards her answer for the Bangkok/Thailand question.

Her expected mean-squared prediction error is

$$\begin{aligned} E(MSE) &= E\left(\phi_T\left((x_{Ts1} - x_{is}) + \ldots + (x_{Tsn_T} - x_{is})\right) + \phi_A\left((x_{As1} - x_{is}) + \ldots + (x_{Asn_A} - x_{is})\right) + x_{it} \ - \mu_t\right)^2 \\ &= E\left(\phi_T \sum_{k=1}^{n_T}(\varepsilon_{Ttk} + \varepsilon_{Tak}) + \phi_A \sum_{k=1}^{n_A}(\varepsilon_{Atk} + \varepsilon_{Aak}) - (n_T\phi_T + n_A\phi_A)(\varepsilon_{it} + \varepsilon_{ia}) + \varepsilon_{it}\right)^2 \\ &= n_T\phi_T^2\left[(\Delta^2_{Tt} + \Delta^2_{Ta}) + (n_T - 1)(\rho_{Tt}\Delta^2_{Tt} + \rho_{Ta}\Delta^2_{Ta})\right] + n_A\phi_A^2[\Delta^2_{At} + \Delta^2_{Aa} + (n_A - 1)(\rho_{At}\Delta^2_{At} + \\ &\quad (1 - n_T\phi_T - n_A\phi_A)^2\alpha\Delta^2_{Tt} + (n_T\phi_T + n_A\phi_A)^2\alpha\Delta^2_{Ta} + 2(1 - n_T\phi_T - n_A\phi_A)n_T\phi_T\alpha\rho_{Tt}\Delta^2_{Tt} \\ &\quad -2(n_T\phi_T + n_A\phi_A)n_T\phi_T\alpha\rho_{Ta}\Delta^2_{Ta} \ , \end{aligned}$$

where ϕ_A is the weight given to American answers for the sum question and ϕ_T is the weight given to other Thai answers for the sum question.

Taking the derivatives with respect to ϕ_A and ϕ_T gives the optimal weights. The optimal weight ratio $\frac{\phi_A}{\phi_T}$ can be expressed as a function of the earlier optimal weight ratios: $\left(\frac{\beta_A}{\beta_T}\right)_{Thai}$ and $\left(\frac{\beta_A}{\beta_T}\right)_{US}$. The first of these is the optimal weight ratio for the Thai questions when subjects observe answers for the Thai questions, and the second is the optimal weight ratio for the US questions when subjects observe information for the US questions. The optimal weight ratio for the Thai question when answers are observed for the sum question only is

$$\left(\frac{\phi_A}{\phi_T}\right)_{Thai} = \frac{\left(\frac{\beta_A}{\beta_T}\right)_{Thai} + \left(\frac{\beta_A}{\beta_T}\right)_{US}\frac{y_A}{y_T} + \frac{cn_T\Delta^2_{Ta}(\rho_{Tt}-\rho_{Ta})(1-\rho_{Ta})}{y_T}}{1 + \frac{y_A}{y_T}\frac{1-\rho_{Tt}}{1-\rho_{Ta}} + \frac{cn_T\Delta^2_{Ta}(\rho_{Ta}-\rho_{Tt})}{y_T}}. \tag{1.28}$$

where

$$\begin{aligned} y_A &= (1+(n_A-1)\rho_{Aa})(1-\rho_{Ta}) \\ y_T &= (1+(n_A-1)\rho_{At})(1-\rho_{Tt}) \end{aligned}$$

If the perceived group bias shares for Thais for the Thailand and US questions, ρ_{Tt} and ρ_{Ta}, are zero, this reduces to the following simple expression

$$\left(\frac{\phi_A}{\phi_T}\right)_{Thai} = \frac{\Delta^2_{Tt} + \Delta^2_{Ta}}{\Delta^2_{At} + \Delta^2_{Aa}} \tag{1.29}$$

This means that

$$(\rho_{Tt})_{perceived} = (\rho_{Ta})_{perceived} = 0 \Rightarrow \left(\frac{\phi_A}{\phi_T}\right)_{Thai} = \left(\frac{\phi_A}{\phi_T}\right)_{US} \tag{1.30}$$

This gives a test for the hypothesis, I_0, that the perceived group bias shares are zero.

$$I_0 : (\rho_{Tt})_{perceived} = (\rho_{Ta})_{perceived} = 0$$

This hypothesis states that subjects ignore group bias for Thais for the Thai and US questions. Notice that I_0 encompasses N_0, the hypothesis that subjects perceive total MSE correctly and ignore group bias, and B_0, the hypothesis that subjects perceive Americans to have lower MSE relative to Thais than they actually do and that subjects ignore group bias. Rejection of I_0

means B_0 must also be rejected.

Under I_0, subjects will choose

$$\left(\frac{\phi_A}{\phi_T}\right)_{Thai} = \left(\frac{\phi_A}{\phi_T}\right)_{US}$$

This test of I_0 is a test of the equality between the weight ratio subjects use to relatively weight American answers for the sum question compared to Thai answers when they update for the Thai question and the analogous ratio when they update for the US question.

The following two regressions give the estimates $\widehat{\phi}_{A,Thai}$, $\widehat{\phi}_{T,Thai}$, $\widehat{\phi}_{A,US}$, and $\widehat{\phi}_{T,US}$ needed to test I_0. The first equation expresses the change in a subject's answer for the Thai question (subscript t) as a function of the distance between the average observed American answer for the sum question and her own answer for the sum question (subscript s) and the distance between the average observed Thai answer for the sum question and her own answer for the sum question.

$$\begin{aligned} y_{iqt} - x_{iqt} &= \phi_A(\overline{x}_{iAqs} - x_{iqs}) + \phi_T(\overline{x}_{iTqs} - x_{iqs}) + \varepsilon_{iqt} && (1.31) \\ y_{iqa} - x_{iqa} &= \phi_A(\overline{x}_{iAqs} - x_{iqs}) + \phi_T(\overline{x}_{iTqs} - x_{iqs}) + \varepsilon_{iqa} && (1.32) \end{aligned}$$

Table 11 reports the results from estimating equations (1.31) and (1.32). Notice that $\left(\frac{\hat{\phi}_A}{\hat{\phi}_T}\right)_{Thai} = \frac{.075}{.057} = 1.32$ and $\left(\frac{\hat{\phi}_A}{\hat{\phi}_T}\right)_{US} = \frac{.225}{.081} = 2.78$. We can reject I_0 at a 10% level ($p = 0.088$). At a 10% level, it is possible to reject that subjects fail to take group bias into account, *regardless of their perceptions American relative accurac compared to Thais.*

In contrast, accounting for group bias would lead to a pattern for the optimal weights similar to that seen in the data for the actual weights. If the perceived group bias share for Thais for the Boston/US questions (ρ_{Ta}) is greater than the perceived group bias share for Thais for the Bangkok/Thailand questions (ρ_{Tt}), then subjects will put a higher relative weight on observed Americans for the US questions than for the Thailand questions.

$$\rho_{Ta} > \rho_{Tt} \Rightarrow \left(\frac{\hat{\phi}_A}{\widehat{\phi}_T}\right)_{Thai} < \left(\frac{\hat{\phi}_A}{\widehat{\phi}_T}\right)_{US}$$

To understand the intuition, consider a subject updating her answer for the Bangkok/Thailand question after observing answers for the sum question. If ρ_{Tt} is high, she should put less weight on Thais relative to Americans for the same reasons seen earlier; group bias means there is less new information in each additional Thai answer. On the other hand, if ρ_{Ta} is high, she should put *higher* weight on observed Thai answers for the sum question. When ρ_{Ta} is high, Thai subjects have a better idea of what other Thai answers about the sum mean for what those observed students believe about the Thai question. For example, if ρ_{Ta} was one, a subject could deduce the observed individual's private signal about the Thai question from her answer to the sum question.

Can the effects of group bias be large enough to replicate the pattern seen in the data? The answer is yes. If subjects applied the estimated actual variance estimates and estimated actual group bias shares, they would choose $\left(\frac{\hat{\phi}_A}{\hat{\phi}_T}\right)_{Thai} = 1.51$ and $\left(\frac{\hat{\phi}_A}{\hat{\phi}_T}\right)_{US} = 2.53$. Figure 3 shows how the optimal weight ratios for the two types of questions vary as a function of ρ_{Ta}, holding the other parameters constant at their estimated values.

The results in the previous section that showed subjects used approximately the optimal weight ratio for the Boston/US and sum questions could have been explained by either subjects appreciating the importance of group bias or by subjects overestimating $\widehat{\frac{\Delta_T^2}{\Delta_A^2}}$ for those questions. By looking at how subjects update separately for the Bangkok/Thailand and Boston/US questions when they observe answers for the sum, it is possible to reject the latter possibility. On the other hand, the actual behavior seen in the data fits closely with subjects appreciating the fact that a bigger share of their error will be attributable to the group for the Boston/US questions than for the Bangkok/Thailand questions.

1.7 Potential gains from changing behavior

Described here are comparisons of the mean-squared error that subjects actually achieve with their initial and final answers and what they could have optimally achieved. Then, I compare how well subjects could have done using different subsets of the information provided to them. Finally, I show that the subjects could primarily improve their performance not by changing how they weigh observed American answers relative to observed Thai answers, but from how they weigh the initial answers they give before observing answers given by other students.

1.7.1 Improvements achieved by subjects

Regressions parallel to (1.13) provide direct estimates of the decreases in MSE that subjects achieve by incorporating the information that they observe. These improvements can be compared to what subjects could have achieved had they used information optimally. The mean-squared distance from the truth achieved by using the optimal weights from regression (1.13) is given by

$$MSD_{(1.13)} = \overline{(Truth_q - \widehat{Truth_q})^2} \; .$$

This optimal MSE can be compared to:

1) The mean-squared distance between the truth and subjects' initial answers, x_{isq}, and

2) The mean-squared distance between the truth and subjects' final answers after observing Thai and American answers, y_{iq}.

The first two rows of Table 12 report these measures. For the Bangkok/Thailand questions, the MSE of subjects' initial answers is 2.2 times greater than the MSE that subjects could have achieved by optimally using all observed answers. Subjects could achieve an improvement of 1.2 by this measure. With their initial answers, subjects attain a MSE that is 5.5 times greater than optimal for the questions about the US. When observing answers for the sum question, subjects attain an MSE that is 1.6 times greater than optimal for the Thai question and 2.0 times greater than optimal for the US question. Significantly smaller gains are attainable when subjects do not observe direct information.

With their actual forecasts, y_{iq}, subjects achieve significant gains by using the information that they observe. Subjects achieve MSEs of 1.8 and 3.1, for the Bangkok/Thailand and Boston/US questions when observing direct information. They achieve MSEs of 1.4 and 1.8 when observing answers for the sum question. When they observe direct information, subjects achieve 32% of the possible improvement over their initial answers for the Bangkok/Thailand questions and 54% of the possible improvement for the Boston/US questions. When they observe answers for the sum question, subjects achieve 33% of the possible improvement over their initial answers for the Bangkok/Thailand questions and 21% of the possible improvement for the Boston/US questions.

1.7.2 Estimating the value of each group's information

Estimating how valuable the American (Thai) information is to a Thai subject entails comparing the MSE that could have been achieved using only the observed American (Thai) answers and a subject's private signal. Each of the following measures is compared to the MSE that could have been achieved by optimally using all observed information:

1) The MSE that could be achieved by optimally using only the observed Thai answers. This can be estimated from the regression equation:

$$Truth_q = \varphi'_s z_{isq} + \varphi'_T \bar{z}_{iTq} + \varepsilon'_{iq}$$

2) The MSE that could be achieved by optimally using only the observed American answers. This can be estimated from the regression equation:

$$Truth_q = \varphi''_s z_{isq} + \varphi''_A \bar{z}_{iAq} + \varepsilon''_{iq}$$

The third row of Table 12 shows the MSEs that subjects could achieve by using just the Thai information and their own initial answers. Relative to the optimal MSE, when observing direct information, subjects could achieve MSE reductions of 83% for the Bangkok/Thailand questions and 50% for the Boston/US questions. When they observe answers for the sum question, the potential improvements attainable by using only the observed Thai answers are 52% and 70% over subjects' initial answers.

The last row of Table 12 shows that, with just the observed American answers, subjects could achieve most of the possible improvement when they observe answers for the Boston/US and sum questions. The potential improvements using only American information are 72% for the Bangkok/Thailand questions and 98% for the Boston/US questions when they observe direct information. They could attain improvements of 93% and 87% when observing answers for the sum question.

Of note is the extra value in American answers compared to Thai answers for the sum question. When subjects see answers for the sum question only, they can achieve about 90% of the improvement possible with the American answers alone. Since Americans and Thais have nearly equal MSEs for the sum question, it is primarily group bias that drives this. The independence of American information makes it relatively more valuable to a Thai than more Thai information.

1.7.3 Gains attainable by reducing overconfidence

Inspection of the estimated optimal weights and the estimated actual weights seems to indicate that the costs of overconfidence are much higher than the costs of overweighing or underweighing observed Thai information relative to observed American information. This can be shown directly by estimating how much could be gained if subjects optimally chose β_s, both in terms of subjects reducing their MSE and increasing their prize money.

To make these calculations, first consider the mean-squared error when the estimated actual weights are used. Equation (1.18) is estimated to obtain $\widehat{y_{iq}}$, $\widehat{\beta_A}$, $\widehat{\beta_T}$, and $\widehat{\beta_s}$.

$$y_{iq} = \beta_A \overline{z}_{iAq} + \beta_T \overline{z}_{iTq} + \beta_s z_{iq} + \varepsilon_{iq}$$

The first row in Table 13 shows the MSE (relative to optimal) that subjects would attain by choosing $\widehat{y_{iq}}$.[12]

To estimate the loss that comes from choosing a non-optimal β_s, equation (1.13) is estimated

$$Truth_q = \varphi_A \overline{x}_{iAq} + \varphi_T \overline{x}_{iTq} + \varphi_s x_{isq} + \varepsilon_{iq} \ ,$$

under the restriction $\varphi_A = \frac{\widehat{\beta_A}}{\widehat{\beta_T}} \varphi_T$. This gives the mean-squared distance that could be achieved if subjects optimally chose φ_s, given the estimated actual relative weighting applied to observed American compared to observed Thai information. Table 13 describes how subjects could improve their MSE by changing how they weigh themselves. Table 14 describes potential improvements in prize money.

The tables show that, when subjects observe answers for the Bangkok/Thailand questions, they can achieve 95% of the possible MSE reduction and 90% of the possible prize money increase by correctly choosing φ_s. When subjects observe answers for the Boston/US questions,

[12] Note that the attained MSE is higher than the second row of Table 12, because $\widehat{y_{iq}}$ does not contain the variation induced by the error term.

they can achieve 99% of the possible MSE reduction and 94% of the possible prize money increase by choosing the optimal self-weight. When subjects see answers for the sum question, the results again show that most of the potential gains come from subjects changing how weigh their original answers. For the Bangkok/Thailand questions, subjects can achieve 97% of the possible MSE reduction and 79% of the possible prize money increase by optimally choosing φ_s. For the Boston/US questions, optimally choosing φ_s achieves 99% of the possible MSE reduction and 70% of the possible prize money increase.

For each of the three types of questions, subjects could improve their performances much more by decreasing their overconfidence than they could by changing how they relatively weigh observed American and Thai answers. Given the nearly optimal relative weights that subjects choose for the US and sum questions, this is not surprising.

1.8 Conclusion

This paper reports the results of an experiment designed to investigate how people in developing countries respond to different kinds of information. For questions that do not specifically refer to Bangkok or Thailand, the experimental subjects apply the optimal relative weights to the American and Thai information they observe. I find that, in achieving this optimality, the Thai students in the experiment show a remarkable sophistication in assessing the relative value of observed American information compared to observed domestic information. Specifically, the results show that the Thai students appreciate the extra value in American information due to its independence. They achieve this cognitive success despite displaying considerable overconfidence.

The experimental results in this paper have important implications for our understanding of how agents conduct inference. Despite potential biases coming from salient ethnic labels, the subjects in the experiment treat the information they observe in a sophisticated way. If people in developing countries generally appreciate the value of observing independent sources of information, ensuring access to a variety of information sources might help to change behaviors affecting agricultural output and public health. However, the results suggest that overconfidence will be a siginificant barrier to individuals appreciating the value of any information they observe.

It is interesting that the subjects fail to use the optimal relative weights only for the questions about Bangkok or Thailand. In this case, they overweigh observed Thais compared to observed Americans. They should put a high weight on observed Thais compared to observed Americans, but they choose an even higher relative weight than they optimally would. Future work could investigate whether it is generally the case that agents overweigh other members of their own group relative to outsiders specifically for tasks that refer to the group's presumed area of expertise.

In addition, future research could apply this experiment to American students, to see if and when they appreciate the value in the independence of Thai information. The experimental results reported here show that the ethnic labels attached to information generally do not interfere with Thai subjects making intelligent decisions in terms of weighing observed American answers versus observed Thai answers. It would be interesting to see if we could come to the same conclusion about American students.

Bibliography

[1] Camerer, Colin and Dan Lovallo (1999). "Overconfidence and Excess Entry: An Experimental Approach," *American Economic Review, 89/1,* 306-318.

[2] Cesarini, David, Orjan Sandewall, and Magnus Johanneson (2003). "Confidence Interval Estimation Tasks and the Economics of Overconfidence," Stockholm School of Economics working paper.

[3] Dearden, Kirk, Lant Pritchett and Jeff Brown (2004). "Learning from Neighbors: Social Learning about Child Feeding During Dysentery Episodes," working paper.

[4] Duflo, Esther, and Emanuel Saez (2003). "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment," *Quarterly Journal of Economics, 118/3,* 815-842.

[5] Duflo, Esther, Michael Kremer and Jonathan Robinson (2003). "Understanding Technology Adoption: Fertilizer in Western Kenya," working paper.

[6] Foster, Andrew and Mark Rosenzweig (1995). "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy, 103*, 1176-1209.

[7] Genest, Christian and Mark J. Schervish (1985). "Modeling Expert Judgments for Bayesian Updating," *Annals of Statistics, 13*, 1198-1212.

[8] Gigerenzer, G., U. Hoffrage, and H. Kleinbolting (1991). "Probabilistic Mental Models: A Brunswikian Theory of Confidence," *Psychological Review, 98*, 506-528.

[9] Griffin, Dale and Amos Tversky (1992). "The Weighting of Evidence and the Determinants of Confidence," *Cognitive Psychology, 24,* 411-435.

[10] Hsieh, Chee-Ruey and Shin-Jong Lin (1997). "Health Information and the Demand for Preventative Care among the Elderly in Taiwan," *Journal of Human Resources, 32,* 308-333.

[11] Keren, G. (1987). "Facing Uncertainty in a Game of Bridge: A Calibration Study," *Organizational Behavior and Human Decision Processes, 39,* 98-114.

[12] Larrick, Richard and Jack Soll (2003). "Intuitions about Combining Opinions: Misappreciation of the Averaging Principle," INSEAD working paper.

[13] Menon, Tanya and Jeffrey Pfeffer (2003), "Valuing Internal vs. External Knowledge: Explaining the Preference for Outsiders," *Management Science.*

[14] Miguel, Edward and Michael Kremer (2003). "Networks, Social Learning, and Technology Adoption: The Case of Deworming Drugs in Kenya," working paper.

[15] Munshi, Kaivan and Jacques Myaux (2002), "Development as a Process of Social Change: An Application to the Fertility Transition," working paper.

[16] Munshi, Kaivan (2004). "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution," *Journal of Development Economics, 73/1,* 185-213.

[17] Tversky, Amos and Daniel Kahneman (1974). "Judgment Under Uncertainty: Heuristics and Biases," *Science, 185,* 1124-1131.

[18] West, Mike (1992). "Modeling Agent Forecast Distributions," *Journal of the Royal Statistical Society, 54*, 553-567.

[19] West, Mike and Jo Crosse (1992). "Modeling Probabilistic Agent Opinion," *Journal of the Royal Statistical Society, 54*, 285-299.

[20] Wright, G.N., et al. (1978). "Cultural Differences in Probabilistic Thinking," *Journal of Cross-Cultural Psychology, 9,* 285-299.

[21] Yates, J. Frank, et al. (2002). "Probability Judgment Across Cultures," in *Heuristics and Biases: The Psychology of Intuitive Judgment,* Gilovich, Griffin, and Kahneman, eds., 271-291.

1 Appendix 1: The questionnaire

1) At the equator, the sun sets 12 hours after it rises on any day. In other places, the day is longer than 12 hours in the summer and shorter than 12 hours in the winter.

Length of longest day in Bangkok		Length of longest day in Boston		Sum
______ hours and ______ minutes	+	______ hours and ______ minutes	=	______ hours and ______ minutes

2)

Highest recorded temperature in Bangkok in 2002		Highest recorded temperature in Boston in 2002		Sum
______ °F	+	______ °F	=	______ °F

3)

From 1961-1990, average number of days <u>per year</u> with recordable precipitation (of any type) in Bangkok		From 1961-1990, average number of days <u>per year</u> with recordable precipitation (of any type) in Boston		Sum
______ days	+	______ days	=	______ days

4)

From 1961-1990, average daily <u>high</u> temperature in January in Bangkok		From 1961-1990, average daily <u>high</u> temperature in January in Boston		Sum
______ °F	+	______ °F	=	______ °F

5) Distance between Bangkok and Boston = ____________ miles

6) The per capita gross national product (GNP) is a measure of the annual mean income that is earned per person in a country. So per capita GNP is the total income of a country divided by the total number of people in that country (including adults and children).

In 2002, per capita GNP of Thailand (in US dollars)		In 2002, per capita GNP of US (in US dollars)		Sum
$ ________	+	$ ________	=	$ ________

7)

2002 population of Thailand		2002 population of US		Sum
________ million people	+	________ million people	=	________ million people

8)

Since January 1, 1960, number of Thai prime ministers		Since January 1, 1960, number of US presidents		Sum
________	+	________	=	________

9)

On October 1, 2003, average price of a gallon of premium gasoline (95 octane) in Bangkok		On October 1, 2003, average price of a gallon of premium gasoline (95 octane) in Boston		Sum
$ __.___	+	$ __.___	=	$ __.___

10) The baht is the currency of Thailand. On January 1, 2003, what was the official Thai baht-to-US dollar exchange rate? (Please give your answer by filling in the blank)

______ baht = $1

11)

In 2002, percentage of Thai women aged 15-24 who were infected with HIV		In 2002, percentage of American women aged 15-24 who were infected with HIV		Sum
________ %	+	________ %	=	________ %

12)

In 2000, percentage of Thai workers who were women (average for the year)		In 2000, percentage of American workers who were women (average for the year)		Sum
________ %	+	________ %	=	________ %

13)

In 2000, percentage of Thais aged 25-29 with at least some university education		In 2000, percentage of Americans aged 25-29 with at least some university education		Sum
________ %	+	________ %	=	________ %

14)

In 2000, percentage of Thai workers whose primary occupation was in agriculture		In 2000, percentage of American workers whose primary occupation was in agriculture		Sum
_______%	+	_______%	=	_______%

15)

In 2000, percentage of Thais who reported they were of Chinese ethnicity		In 2000, percentage of Americans who reported they were of African ethnicity		Sum
_______%	+	_______%	=	_______%

2 Appendix 2: Proof for Theorem 1

Theorem 1 *Equation (14) is the MLE for $\widehat{\rho}_j$ under the model.*

Proof. Consider group j (either A or T). For a given question q, it is clear that

$$\widehat{\sigma^2_{jq}} = \frac{1}{N}\sum_{i=1}^{N}(x_{ijq} - Truth_q)^2 \tag{A1}$$

is the MLE for the true mean-squared distance of group j answers from the truth, σ^2_{Tq}. In the model, the sample variance times $\frac{N-1}{N}$ is $(1-\rho)\sigma^2_{Tq}$. It is thus clear that

$$(1-\widehat{\rho_q})\sigma^2_{jq} = \frac{1}{N}\sum_{i=1}^{N}(x_{ijq} - \overline{x}_{jq})^2 \tag{A2}$$

is the MLE for the sample variance, where $\overline{x}_{Tq}$ is the average of the Thai answers for question q, and ρ_q contains the subscript to denote that the estimate of ρ is based on question q. Since the MLE of a function is the function of the MLEs, we can separate (A2), and substitution of (A1) into (A2) gives

$$1-\widehat{\rho}_q = \frac{\sum_{i=1}^{N}(x_{ijq} - \overline{x}_{jq})^2}{\sum_{i=1}^{N}(x_{ijq} - Truth_q)^2},$$

which leads to

$$\widehat{\rho}_q = \frac{\sum_{i=1}^{N}(\overline{x}_{jq} - Truth_q)^2}{\sum_{i=1}^{N}(x_{ijq} - Truth_q)^2} = \frac{\widehat{\alpha^2_{jq}}}{\widehat{\sigma^2_{jq}}} \ .$$

Since the questions are assumed to be independent, the MLE for ρ is just the average of the estimates provided by all the Q questions:

$$\widehat{\rho} = \frac{1}{Q}\sum_{q=1}^{Q}\widehat{\rho}_q = \frac{1}{Q}\sum_{q=1}^{Q}\frac{\widehat{c^2_{jq}}}{\widehat{\sigma^2_{jq}}} \ ,$$

as was to be shown. ■

From 1961-1990, average daily high temperature in January in Bangkok		From 1961-1990, average daily high temperature in January in Boston		Sum
_____ °C	+	_____ °C	=	_____ °C

Answers given by Thai students

1. _23_ °C
2. _20_ °C

Answers given by American students

1. _2_ °C

Figure 1: Sample of information shown to subjects

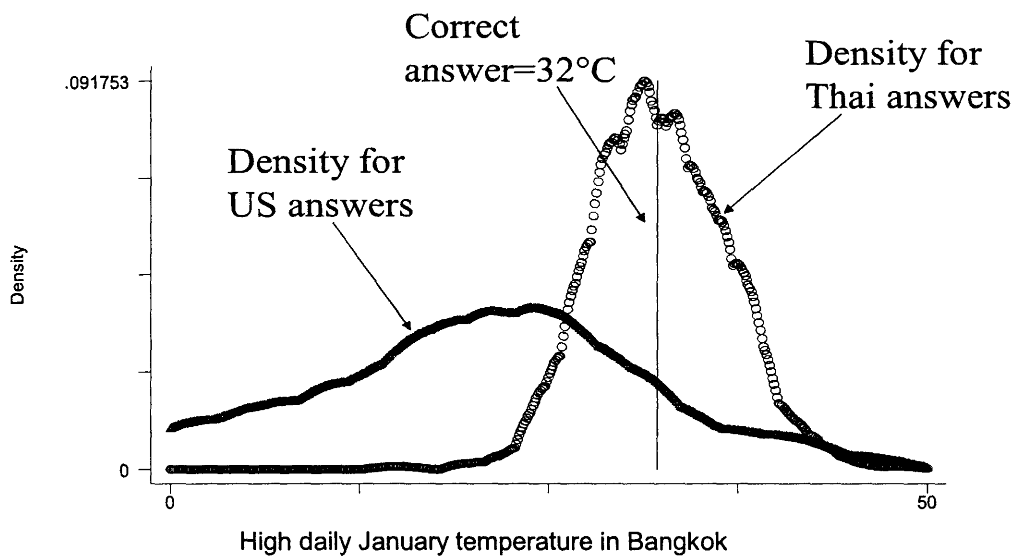

Figure 2A: Kernel density estimates for high daily January temperature in Bangkok (in °C)

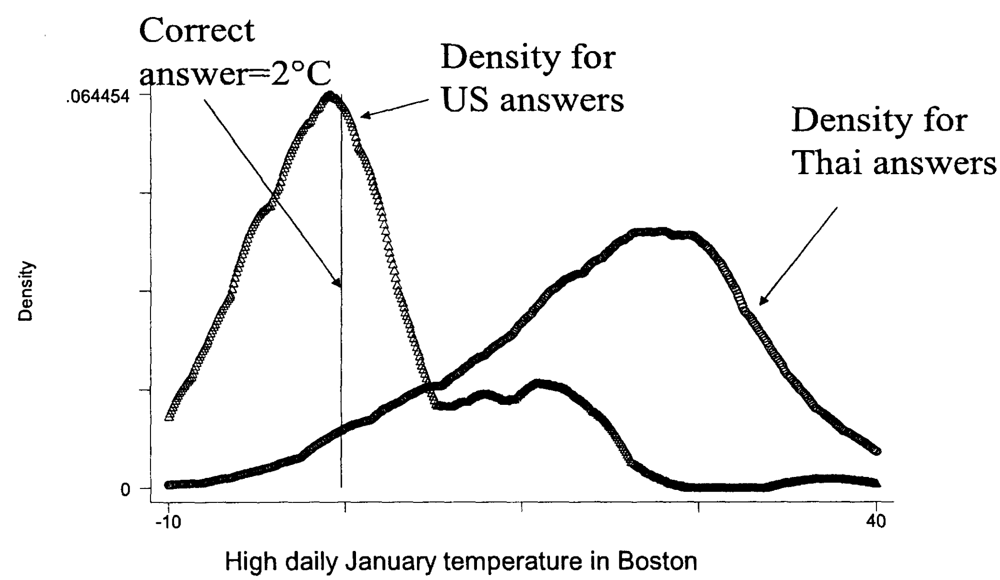

Figure 2B: Kernel density estimates for high daily January temperature in Boston (in °C)

Notes: (1) The Epanechnikov kernel is used with optimal bandwidth selection.
(2) Density estimates are estimated over different ranges in the Figures 2A and 2B.

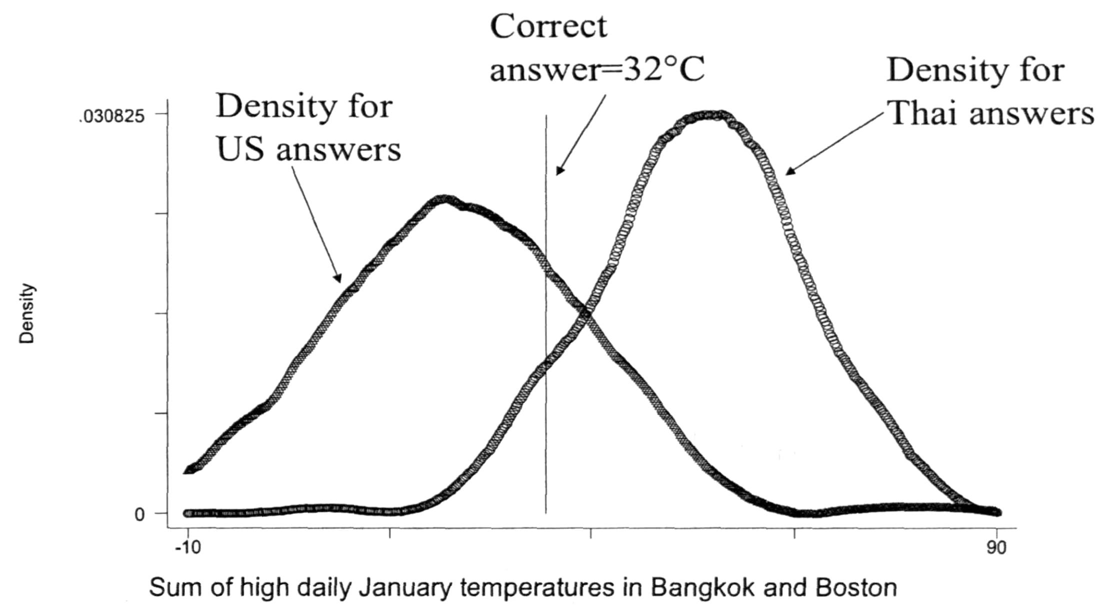

Figure 2C: Kernel density estimates for sum of high daily January temperatures (in °C)

Note: (1) The Epanechnikov kernel is used with optimal bandwidth selection.

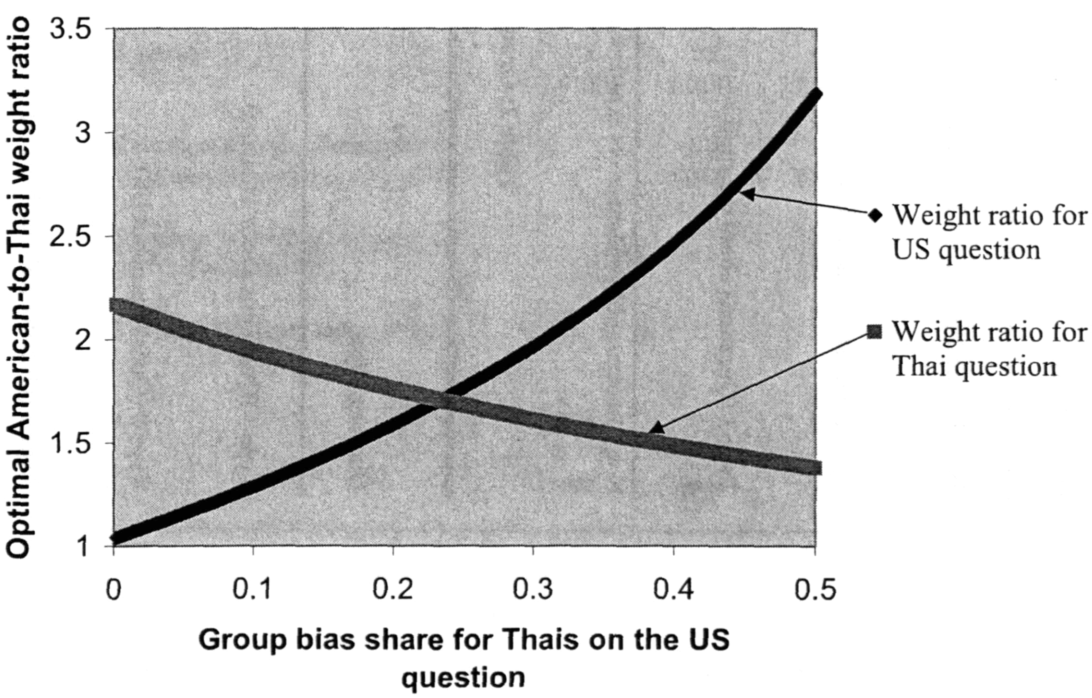

Figure 3: The American-to-Thai weight ratio as a function of the Thai group bias share for the US questions

Table 1: Relative accuracy of Americans and Thais

Question type	$\frac{MSE_{Thai}}{MSE_{Thai} + MSE_{US}}$ (1)	$\frac{MSE_{Thai}}{MSE_{US}}$ (2)
Type 1 (Questions about Thailand)	.341 (.008)	.517 (.018)
Type 2 (Questions about US)	.755 (.013)	3.086 (.216)
Type 3 (Questions about the sum)	.565 (.01)	1.299 (.052)

Note: Bootstrapped standard errors in parentheses

Table 2: Group effects for American and for Thais

Question type	Estimated Thai group bias share (1)	Estimated American group bias share (2)
Type 1 (Questions about Thailand)	.234 (.066)	.307 (.056)
Type 2 (Questions about US)	.362 (.088)	.227 (.063)
Type 3 (Questions about the sum)	.336 (.081)	.277 (.062)

Note: Bootstrapped standard errors in parentheses

Table 3A: Weights used when both types of information are seen (Thailand questions)

Dependent variable: Subjects' final answers for the Thailand questions

Regressor	Questions about Thailand			
	(1)	(2)	(3)	(4)
Own prior ($\beta_{S,1}$)	.654 (.016)	.655 (.016)	.651 (.016)	.661 (.018)
Own prior*number of observed American answers ($\beta_{S,2}$)		.002 (.019)	.008 (.019)	.004 (.019)
Own prior*number of observed Thai answers ($\beta_{S,3}$)			-.038 (.019)	-.04 (.019)
Own prior*accuracy index ($\beta_{S,4}$)				-.142 (.084)
Thai average ($\beta_{T,1}$)	.242 (.02)	.245 (.021)	.245 (.025)	.243 (.026)
Thai average*number of observed American answers ($\beta_{T,2}$)		.002 (.024)	.001 (.024)	.011 (.025)
Thai average*number of observed Thai answers ($\beta_{T,3}$)			.001 (.025)	-.002 (.025)
Thai average*accuracy index ($\beta_{T,4}$)				.036 (.096)
US average ($\beta_{A,1}$)	.058 (.012)	.05 (.014)	.039 (.015)	.063 (.016)
US average*number of observed American answers ($\beta_{A,2}$)		-.016 (.017)	0 (.019)	-.003 (.019)
US average*number of observed Thai answers ($\beta_{A,3}$)			-.036 (.018)	-.026 (.018)
US average*accuracy index ($\beta_{A,4}$)				.232 (.055)
R^2	.979	.979	.979	.98
N	1008	986	986	986

Notes:

(1) Regression standard errors are in parentheses.

(2) R-squared values are calculated using the data standardized by the standard deviations only, and not the means.

(3) The regression in column (1) includes the data for questions 4 and 8 where subjects saw either 0, 5, 10, or 20 American and Thai answers.

(4) Regressions include dummies for the three question categories (meteorology, economic/political, and social/cultural).

Table 3B: Weights used when both types of information are seen (US questions)

Dependent variable: Subjects' final answers for US questions

Regressor	Questions about US (1)	(2)	(3)	(4)
Own prior ($\beta_{S,1}$)	.461	.477	.482	.493
	(.02)	(.02)	(.021)	(.021)
Own prior*number of observed American answers ($\beta_{S,2}$)		.002	0	.012
		(.024)	(.024)	(.024)
Own prior*number of observed Thai answers ($\beta_{S,3}$)			.04	.044
			(.024)	(.024)
Own prior*accuracy index ($\beta_{S,4}$)				-.271
				(.128)
Thai average ($\beta_{T,1}$)	.092	.088	.056	.065
	(.03)	(.03)	(.036)	(.036)
Thai average*number of observed American answers ($\beta_{T,2}$)		-.016	-.005	.002
		(.034)	(.034)	(.037)
Thai average*number of observed Thai answers ($\beta_{T,3}$)			-.051	-.037
			(.037)	(.041)
Thai average*accuracy index ($\beta_{T,4}$)				-.116
				(.219)
US average ($\beta_{A,1}$)	.466	.484	.478	.458
	(.019)	(.02)	(.02)	(.021)
US average*number of observed American answers ($\beta_{A,2}$)		.075	.089	.07
		(.021)	(.023)	(.023)
US average*number of observed Thai answers ($\beta_{A,3}$)			-.046	-.053
			(.024)	(.024)
US average*accuracy index ($\beta_{A,4}$)				.387
				(.124)
R^2	.924	.926	.926	.927
N	1053	1032	1032	1032

Notes:
(1) Regression standard errors are in parentheses.
(2) R-squared values are calculated using the data standardized by the standard deviations only, and not the means.
(3) The regression in column (1) includes the data for questions 4 and 8 where subjects saw either 0, 5, 10, or 20 American and Thai answers.
(4) Regressions include dummies for the three question categories (meteorology, economic/political, and social/cultural).

Table 3C: Weights used when both types of information are seen (sum questions)

Dependent variable: Subjects' final answers for sum questions

Regressor	Questions about sum (1)	(2)	(3)	(4)
Own prior $(\beta_{S,1})$	.733 (.019)	.733 (.02)	.733 (.02)	.731 (.02)
Own prior*number of observed American answers $(\beta_{S,2})$		.001 (.024)	-.001 (.024)	-.001 (.024)
Own prior*number of observed Thai answers $(\beta_{S,3})$			-.002 (.026)	.001 (.026)
Own prior*accuracy index $(\beta_{S,4})$				.261 (.13)
Thai average $(\beta_{T,1})$	.069 (.02)	.07 (.026)	.063 (.029)	.076 (.032)
Thai average*number of observed American answers $(\beta_{T,2})$		.004 (.03)	.009 (.031)	.006 (.031)
Thai average*number of observed Thai answers $(\beta_{T,3})$			-.024 (.037)	-.016 (.037)
Thai average*accuracy index $(\beta_{T,4})$				-.151 (.204)
US average $(\beta_{A,1})$	.167 (.024)	.096 (.037)	.091 (.037)	.084 (.037)
US average*number of observed American answers $(\beta_{A,2})$		-.068 (.031)	-.067 (.031)	-.06 (.032)
US average*number of observed Thai answers $(\beta_{A,3})$			-.023 (.024)	-.011 (.025)
US average*accuracy index $(\beta_{A,4})$				.353 (.138)
R^2	.961	.963	.963	.963
N	557	548	548	548

Notes:

(1) Regression standard errors are in parentheses.

(2) R-squared values are calculated using the data standardized by the standard deviations only, and not the means.

(3) The regression in column (1) includes the data for questions 4 and 8 where subjects saw either 0, 5, 10, or 20 American and Thai answers.

(4) Regressions include dummies for the three question categories (meteorology, economic/political, and social/cultural).

Table 4A: MSE minimizing behavior when both types of answers are seen

Dependent variable: Correct answer for Thailand questions

Regressor	Questions about Thailand			
	(1)	(2)	(3)	(4)
Own prior $(\beta_{S,1})$	.295	.293	.269	.286
	(.021)	(.021)	(.021)	(.019)
Own prior*number of observed American answers $(\beta_{S,2})$		.035	.053	.033
		(.024)	(.024)	(.021)
Own prior*number of observed Thai answers $(\beta_{S,3})$			-.113	-.115
			(.024)	(.02)
Own prior*accuracy index $(\beta_{S,4})$				-.902
				(.086)
Thai average $(\beta_{T,1})$	.458	.435	.497	.403
	(.023)	(.023)	(.025)	(.023)
Thai average*number of observed American answers $(\beta_{T,2})$		-.145	-.187	-.151
		(.029)	(.03)	(.026)
Thai average*number of observed Thai answers $(\beta_{T,3})$			.183	.145
			(.029)	(.025)
Thai average*accuracy index $(\beta_{T,4})$				-.228
				(.092)
US average $(\beta_{A,1})$	.247	.272	.234	.311
	(.018)	(.019)	(.02)	(.018)
US average*number of observed American answers $(\beta_{A,2})$		.11	.134	.118
		(.024)	(.026)	(.022)
US average*number of observed Thai answers $(\beta_{A,3})$			-.07	-.03
			(.025)	(.022)
US average*accuracy index $(\beta_{A,4})$				1.13
				(.065)
N	1008	986	986	986

Notes:
(1) Regression standard errors are in parentheses.
(2) The regression in column (1) includes the data for questions 4 and 8 where subjects saw either 0, 5, 10, or 20 American and Thai answers.
(3) Regressions include dummies for the three question categories (meteorology, economic/political, and social/cultural).
(4) Regressions are esimtated under the constraints:

$$\beta_{S,1} + \beta_{T,1} + \beta_{A,1} = 1 \text{ and } \beta_{S,j} + \beta_{T,j} + \beta_{A,j} = 0 \text{ for } j > 1$$

Table 4B: MSE minimizing behavior when both types of answers are seen

Dependent variable: Correct answer for US questions

Regressor	Questions about US			
	(1)	(2)	(3)	(4)
Own prior $(\beta_{S,1})$	.063	.056	.04	.083
	(.017)	(.017)	(.018)	(.014)
Own prior*number of observed American answers $(\beta_{S,2})$		-.056	-.048	-.029
		(.019)	(.019)	(.016)
Own prior*number of observed Thai answers $(\beta_{S,3})$			-.08	-.095
			(.02)	(.016)
Own prior*accuracy index $(\beta_{S,4})$				-.755
				(.083)
Thai average $(\beta_{T,1})$	.115	.075	.089	.153
	(.02)	(.02)	(.021)	(.017)
Thai average*number of observed American answers $(\beta_{T,2})$		-.137	-.135	-.056
		(.023)	(.023)	(.019)
Thai average*number of observed Thai answers $(\beta_{T,3})$			.036	.119
			(.022)	(.018)
Thai average*accuracy index $(\beta_{T,4})$				-1.402
				(.1)
US average $(\beta_{A,1})$	.822	.869	.871	.764
	(.016)	(.017)	(.017)	(.014)
US average*number of observed American answers $(\beta_{A,2})$		.193	.183	.085
		(.019)	(.019)	(.015)
US average*number of observed Thai answers $(\beta_{A,3})$			.043	-.024
			(.019)	(.015)
US average*accuracy index $(\beta_{A,4})$				2.157
				(.082)
N	1053	1032	1032	1032

Notes:
(1) Regression standard errors are in parentheses.
(2) The regression in column (1) includes the data for questions 4 and 8 where subjects saw either 0, 5, 10, or 20 American and Thai answers.
(3) Regressions include dummies for the three question categories (meteorology, economic/political, and social/cultural).
(4) Regressions are esimtated under the constraints:

$$\beta_{S,1} + \beta_{T,1} + \beta_{A,1} = 1 \text{ and } \beta_{S,j} + \beta_{T,j} + \beta_{A,j} = 0 \text{ for } j > 1$$

Table 4C: MSE minimizing behavior when both types of answers are seen

Dependent variable: Correct answer for sum questions

Regressor	Questions about sum			
	(1)	(2)	(3)	(4)
Own prior ($\beta_{S,1}$)	.18	.148	.152	.14
	(.026)	(.028)	(.03)	(.022)
Own prior*number of observed American answers ($\beta_{S,2}$)		-.117	-.118	-.121
		(.034)	(.034)	(.025)
Own prior*number of observed Thai answers ($\beta_{S,3}$)			.018	-.045
			(.036)	(.027)
Own prior*accuracy index ($\beta_{S,4}$)				-.525
				(.133)
Thai average ($\beta_{T,1}$)	.25	.249	.247	.396
	(.024)	(.032)	(.035)	(.027)
Thai average*number of observed American answers ($\beta_{T,2}$)		.036	.037	.022
		(.035)	(.035)	(.026)
Thai average*number of observed Thai answers ($\beta_{T,3}$)			-.006	0
			(.036)	(.027)
Thai average*accuracy index ($\beta_{T,4}$)				-2.314
				(.154)
US average ($\beta_{A,1}$)	.57	.603	.602	.464
	(.026)	(.037)	(.038)	(.03)
US average*number of observed American answers ($\beta_{A,2}$)		.081	.081	.098
		(.033)	(.033)	(.025)
US average*number of observed Thai answers ($\beta_{A,3}$)			-.012	.045
			(.031)	(.023)
US average*accuracy index ($\beta_{A,4}$)				2.838
				(.14)
N	557	548	548	548

Notes:
(1) Regression standard errors are in parentheses.
(2) The regression in column (1) includes the data for questions 4 and 8 where subjects saw either 0, 5, 10, or 20 American and Thai answers.
(3) Regressions include dummies for the three question categories (meteorology, economic/political, and social/cultural).
(4) Regressions are esimtated under the constraints:

$$\beta_{S,1} + \beta_{T,1} + \beta_{A,1} = 1 \text{ and } \beta_{S,j} + \beta_{T,j} + \beta_{A,j} = 0 \text{ for } j > 1$$

Table 5A: Estimated actual weights when both types of information are seen

Actual weight	Thailand questions	US questions	Sum questions
	(1)	(2)	(3)
Own prior	.654	.461	.733
	(.016)	(.02)	(.019)
Thai average	.242	.092	.069
	(.02)	(.03)	(.02)
US average	.058	.466	.167
	(.012)	(.019)	(.024)
N	1008	1053	557

Note:

(1) Regression standard errors are in parentheses.

(2) These estimates come from the regressions in column 1 in Tables 3A, 3B, and 3C.

Table 5B: Optimal weights estimated from reduced form regression

Optimal weight	Questions about Thailand	Questions about US	Questions about sum
	(1)	(2)	(3)
Own prior	.295	.063	.18
	(.021)	(.017)	(.026)
Thai average	.458	.115	.25
	(.023)	(.02)	(.024)
American average	.247	.822	.57
	(.018)	(.016)	(.026)

Notes:
(1) Regression standard errors are in parentheses.
(2) These estimates come from the regressions in column 1 in Tables 4A, 4B, and 4C.

Table 5C: Optimal weights estimated from econometric model (given $c = 1$)

Optimal weight	Questions about Thailand	Questions about US	Questions about sum
	(1)	(2)	(3)
Own prior	0.258	0.100	0.155
	(0.007)	(0.008)	(0.009)
Thai average	0.459	0.174	0.333
	(0.014)	(0.015)	(0.020)
American average	0.283	0.726	0.512
	(0.021)	(0.024)	(0.028)

Notes:
(1) Bootstrapped standard errors are in parentheses.
(2) These estimates come from the parameter estimates described in Tables 1, 2, and 3.

Table 6: Effect of large samples on a subject's choice to update

A=main group (questions 4 and 8), B=big sample treatment group

	A	B	A	B	A	B
	Questions about Thailand		Questions about US		Questions about sum	
	(1)	(2)	(3)	(4)	(5)	(6)
Share of movers	.384	.32	.55	.45	.644	.54
N	500	100	500	100	500	100

Notes:
(1) Subjects in the big sample treatment group saw 0, 5, 10, or 20 answers of each type.
(2) Subjects in the main group saw 0, 1, 2, or 3 answers of each type.

Table 7: Effect of anchoring on a subject's MSE

Dependent variable: Squared distance from the correct answer

	Questions about Thailand		Questions about US		Questions about sum	
	(1)	(2)	(3)	(4)	(5)	(6)
All final answers						
Dummy for group that did not provide private beliefs	-0.029 (0.093)	-0.039 (0.093)	-0.071 (0.220)	-0.049 (0.221)	-0.063 (0.148)	-0.055 (0.148)
Version dummies?	N	Y	N	Y	N	Y
N	4235	4235	4196	4196	4192	4192
Only final answers for those subjects who see both kinds of answers for the given question type						
Dummy for group that did not provide private beliefs	0.05 (0.171)	0.036 (0.169)	-0.137 (0.287)	-0.235 (0.280)	0.027 (0.374)	-0.001 (0.373)
Version dummies?	N	Y	N	Y	N	Y
N	1132	1132	1185	1185	626	626

Note:
(1) Regression standard errors are in parentheses.

Table 8: Effect of anchoring on distance from observed information

Independent variable: Dummy for anchoring treatment group

Dependent variable	All observations	Only Thai answers seen	Only American answers seen	Average squared distance for subjects in main group
	(1)	(2)	(3)	(4)
Questions about Thailand				
Squared distance from average of observed Thai answers	-.014 (.021) N=1561	-.014 (.041) N=429		.832 (1.778) N=1358
Squared distance from average of observed American answers	.043 (.079) N=1510		-.043 (.129) N=378	1.727 (2.584) N=1314
Questions about US				
Squared distance from average of observed Thai answers	-.051 (.065) N=1563	-.161 (.087) N=378		1.798 (3.23) N=1360
Squared distance from average of observed American answers	-.046 (.022) N=1566		.013 (.013) N=381	1.378 (2.945) N=1360
Questions about sum				
Squared distance from average of observed Thai answers	-.026 (.049) N=1555	-.123 (.079) N=346		1.75 (3.585) N=1353
Squared distance from average of observed American answers	-.049 (.07) N=1644		-.184 (.097) N=435	1.781 (3.314) N=1433

Note:

(1) For columns (1), (2) and (3) robust regression standard errors are reported in parentheses.
For column (4), the standard deviation is in parentheses.

Table 9: NLS estimate of overconfidence and perceived correlation

Non-linear regressions with $\frac{\Delta_T^2}{\Delta_A^2}$ estimated as an endogenous parameter.

Parameter	Thailand questions	US questions	Sum questions
	(1)	(2)	(3)
$c(n_A + n_T = 2)$	.657	.465	.359
	(.17)	(.339)	(.379)
$c(n_A + n_T = 3)$	.233	.308	.16
	(.145)	(.234)	(.184)
$c(n_A + n_T = 4)$	.582	.412	.244
	(.207)	(.284)	(.277)
$c(n_A + n_T = 5)$	.484	.301	.211
	(.231)	(.205)	(.241)
$c(n_A + n_T = 6)$	.387	.189	.11
	(.225)	(.129)	(.14)
$\rho_A = \rho_T$	.526	.722	.573
	(.314)	(.277)	(.484)
$\frac{\Delta_T^2}{\Delta_A^2}$	.168	6.161	1.236
	(.045)	(1.871)	(.311)
R^2	.98	.925	.961
N	986	1032	548

Notes: (1) Non-linear regression standard errors are in parentheses.
(2) Regressions are estimated under the assumption that the perceived Thai and American correlation coefficients are equal.

Table 10: Comparing the weight ratios

Estimated American-to-Thai weight ratio

	Thailand questions	US questions	Sum questions
	(1)	(2)	(3)
Actual weight ratio	.240	5.065	2.420
90% confidence interval	(.150,.334)	(3.420,11.385)	(1.601,4.501)
95% confidence interval	(.139,.351)	(3.118,14.232)	(1.461,5.577)
Naive weight ratio	.517	3.086	1.299
	(.477,.554)	(2.737,3.577)	(1.205,1.406)
	p=0.000	*p=0.058*	*p=0.017*
Optimal weight ratios for different overconfidence levels			
$c = 1$	.571	4.423	1.843
	(.475,.694)	(3.467,5.63)	(1.466,2.332)
	p=0.001	*p=0.610*	*p=0.367*
$c = 0.75$	.576	4.581	1.913
	(.464,.727)	(3.504,6.29)	(1.469,2.524)
	p=0.001	*p=0.750*	*p=0.457*
$c = 0.5$	.606	4.897	2.048
	(.476,.759)	(3.62,6.938)	(1.549,2.877)
	p=0.000	*p=0.888*	*p=0.600*
$c = 0.25$	.616	5.125	2.135
	(.468,.798)	(3.662,7.592)	(1.525,2.991)
	p=0.000	*p=0.976*	*p=0.709*
Estimated $c(n_A + n_T)$	.603	5.043	2.168
	(.476,.773)	(3.635,7.345)	(1.597,3.136)
	p=0.002	*p=0.963*	*p=0.721*

Notes: (1) Bootstrapped 95% confidence intervals are in parentheses.

(2) The bootstrap for the actual weights accounts for correlation in the coefficient estimates.

(3) p-values compare the given weight ratio to the actual weight ratio.

Table 11: Subject behavior for the Thai and US questions (see answers for the sum)

Dependent variables : (1) Final answer for Thai question
(2) Final answer for US question

Regression weights	Thailand questions (1)	US questions (2)
ϕ_T = Distance between observed Thai average and initial answer (*for sum question*)	.057 (.013)	.081 (.019)
ϕ_A = Distance between observed American average and initial answer (*for sum question*)	.075 (.015)	.225 (.022)
p-value for $H_0 : \left(\frac{\phi_A}{\phi_T}\right)_{Thai} = \left(\frac{\phi_A}{\phi_T}\right)_{US}$	0.088	
N	548	544

Notes:
(1) Regression standard errors are in parentheses.
(2) Regressions include dummies for the three question categories (meteorology, economic/political, and social/cultural).

Table 12: Comparing MSE to that achieved by using all information

Efficiency measure = $\dfrac{x_i}{\textit{Optimal MSE attainable with all observed answers}}$

	Subjects see direct information		Subjects see answers for sum question	
Relative efficiency	Questions about Thailand (1)	Questions about US (2)	Questions about Thailand (3)	Questions about US (4)
x_1 = MSE achieved with initial answer (before observing information)	2.228	5.505	1.595	2.029
x_2 = MSE achieved with final answer (after observing information)	1.838	3.082	1.398	1.816
x_3 = Optimal MSE possible with just initial answer and observed Thai average	1.212	3.272	1.283	1.309
x_4 = Optimal MSE possible with just initial answer and observed American average	1.346	1.104	1.044	1.137

Table 13: Possible gains from changing self-weight (in MSE)

$$\text{Efficiency measure} = \frac{x_i}{\textit{Optimal MSE attainable with all observed answers}}$$

Relative efficiency	Subjects see direct information		Subjects see answers for sum question	
	Questions about Thailand	Questions about US	Questions about Thailand	Questions about US
	(1)	(2)	(3)	(4)
MSE achieved using:				
x_1 = Prediction with estimated weights used by subjects	1.718	1.92	1.696	1.888
x_2 = Prediction with estimated Thai-to-American ratio and optimal self weight	1.039	1.01	1.022	1.011
Fraction of total possible MSE improvement from choosing optimal self-weight	0.946	0.989	0.968	0.988

Notes:

(1) Prediction with estimated weights refers to the MSE achieved by using $\hat{y}_{iq} = \hat{\beta}_A \bar{x}_{iAq} + \hat{\beta}_T \bar{x}_{iTq} + \hat{\beta}_S x_{isq}$ where the coefficients are the estimated actual weights that subjects use.

(2) Prediction with estimated Thai-to-American ratio and the optimal self-weight refers to the MSE achieved by using the estimated $\frac{\hat{\beta}_A}{\hat{\beta}_T}$ and the optimal self weight given $\frac{\hat{\beta}_A}{\hat{\beta}_T}$ and constraining the weights to sum to one.

Table 14: Possible gains from changing self-weight (in prize money)

Efficiency measure = $\dfrac{x_i}{\textit{Prize money attainable with all observed answers}}$

Relative efficiency	Subjects see direct information: Questions about Thailand	Subjects see direct information: Questions about US	Subjects see answers for sum question: Questions about Thailand	Subjects see answers for sum question: Questions about US
	(1)	(2)	(3)	
Prize money achieved using:				
x_1 = Prediction with estimated weights used by subjects	0.587	0.629	0.611	0.804
x_2= Prediction with estimated Thai-to-American ratio and optimal self weight	0.957	0.977	0.917	0.941
Fraction of total possible prize money increase from choosing optimal self-weight	0.896	0.938	0.787	0.699

Notes:
(1) Prediction with estimated weights refers to the MSE achieved by using $\hat{y}_{iq} = \hat{\beta}_A \bar{x}_{iAq} + \hat{\beta}_T \bar{x}_{iTq} + \hat{\beta}_s x_{isq}$ where the coefficients are the estimated actual weights that subjects use.
(2) Prediction with estimated Thai-to-American ratio and the optimal self-weight refers to the MSE achieved by using the estimated $\frac{\hat{\beta}_A}{\hat{\beta}_T}$ and the optimal self weight given $\frac{\hat{\beta}_A}{\hat{\beta}_T}$ and constraining the weights to sum to one.

Chapter 2

Is There an Information Endowment Effect?

Summary 2 *This paper reports the results of an experiment in which subjects were endowed with information that could help them to answer a series of general-knowledge questions. Experimental evidence suggests that there is an endowment effect for goods. On average, an individual who is endowed with a coffee mug or chocolate bar demands a much higher price to sell than an unendowed person is willing to pay to acquire the same good. This study shows that a similar phenomenon does not exist when the endowment consists of information rather than goods. The results suggest that the endowment effect operates primarily on preferences as opposed to judgment. Subjects in the experiment appear not to open separate mental accounts for monetary rewards and the endowed information.*

2.1 Introduction

Experimental evidence indicates that, for a range of consumption goods, individuals generally demand a higher price for selling a good with which they are endowed than they would be willing to pay to acquire the same good (Knetsch (1989), Kahneman, Knetsch, and Thaler (1990), List (2003)). This behavioral anomaly has been termed the endowment effect (Thaler (1980)). The implications of consumption endowment effects are far-reaching. For example, endowment effects cause the Coase Theorem to no longer hold, invalidating neoclassical welfare economics (Kahneman (1979)). Given the profound implications of endowment effects for goods, it is important to understand what causes endowment to affect behavior. It appears that agents open separate accounts for different goods and for money, causing them to demand higher selling prices and to be reluctant to trade endowed goods for unendowed goods. This paper reports the results of an experiment that endows subjects with items that are valuable but cannot be consumed. The experimental design makes it possible to see whether the endowment effect causes agents to open a separate account when the endowment consists of information rather than goods.

The endowment effect is a specific example of the more general phenomenon of loss aversion. Empirical examples of loss aversion include an unwillingness of real estate and stock market investors to sell assets for less than they paid for them (Genesove and Mayer (2001), Odean (1998)). In this case, the loss aversion reflects the fact that agents maintain separate mental accounts for different assets...

In the experiment, students at Loyola Marymount University first answered general knowledge questions about past temperature and recent gas prices in Bangkok and Boston. In separate research (Healy (2004)), I asked the same questions to 130 students at Thammasat University (a top-ranked university in Bangkok) and 116 MIT students. Information endowments were sealed envelopes containing summary statistics about how the Thammasat and MIT students had answered the questions. In one experimental treatment, subjects were endowed either with summary statistics about MIT students' opinions for the temperature in Boston (question 1) or gas prices in Boston (question 2). Subjects who were endowed with information pertaining to question 1 (question 2) were given the opportunity to trade for information pertaining to question 2 (question 1), before they observed the information contained in an envelope. After trading, subjects observed the information in their possession and they were given the opportunity to revise their answers. The trading data provides evidence against the presence of an information endowment effect. Of the 84 subjects, 43 (51%) traded. This is a stark contrast with previous research that has looked at goods such as chocolate bars or mugs, which find that 85%-90% of subjects choose to keep their endowed good.

To provide additional evidence about the presence or absence of information endowment effects, the same subjects also participated in a second treatment in which they were given the opportunity to buy and sell information pertaining to the questions about Bangkok. Subjects were either endowed with summary statistics about Thammasat students' opinions about the temperature in Bangkok (question 3) or gas prices in Thailand (question 4). The experimental design elicited from subjects endowed with information about question 3 (question 4) the price at which they were willing to sell their endowment and the price at which they were willing to buy information about question 4 (question 3). This design allows me to see if subjects put a higher value on information because it is part of their endowment.[1] In keeping with the trading results, the data cannot reject the hypothesis that a subject's valuations of information are not affected by whether that information is part of her endowment.

The rest of this paper is organized as follows. In Section II, I describe the first part of the experimental design, which involved a simple trading experiment, and the results. Section III describes the experimental design and results relating to the treatments that elicited the monetary values that subjects attached to endowed information relative to other information. In Section IV, I discuss the results' implications for our understanding of consumption endowment effects and describe evidence confirming the efficiency of the information market in the experiment. Section V concludes.

2.2 Experimental Design I

2.2.1 The questions

To test for an information endowment effect, I follow a design similar to that employed in papers that have tested for a consumption endowment effect (Knetsch (1989), Kahneman, Knetsch, and Thaler (1990), and List (2003)). These papers test for an endowment effect by employing two treatments. Subjects are endowed with a good in one treatment and a different good in the other. My design differs from previous efforts in that subjects are endowed with information that can help them answer a series of questions.

To start the experiment, each subject took a seat at a study carrel that contained a sealed

[1] An endowment effect could arise either because subjects undervalue goods or information that is not part of the endowment or because the endowment is overvalued. Kahneman, Knetsch, and Thaler (1990) and List (2003) present evidence suggesting that the endowment effect arises primarily through inflated WTA values rather than depressed WTP values.

manila envelope that they were instructed not to open.[2] Like List's (2003) procedure when providing goods endowment, the subjects were given the manila envelopes at the start of the experiment, so as to maximize the time that subjects possessed their endowments before making decisions. Subjects were not informed of the envelopes' contents. Each of the six experimental sessions had fourteen students. Subjects were told that they would be answering some general-knowledge questions about Boston and Bangkok and that they would receive payment for each question on which their answer was among the five closest to the correct answer. They then answered the following four questions:

Question 1: From 1961-90, what was the average high daily temperature in Boston?

Question 2: On October 1, 2003, what was the average price of a gallon of premium gasoline in Boston?

Question 3: From 1961-90, what was the average high daily temperature in Bangkok?

Question 4: On October 1, 2003, what was the average price of a gallon of premium gasoline in Bangkok?

The subjects answered each question without knowing the contents of the manila envelopes that they had been given. Their answers were collected before proceeding to the two kinds of treatments: 1) treatments that involved trading of information, and 2) treatments that involved auctions where subjects had the opportunity to buy and sell information.

2.2.2 Information trading

After handing in their answers to the four questions, subjects were told that their manila envelopes contained information that could help them to answer the questions. Each envelope contained two smaller sealed envelopes, one that pertained to a question about Boston and another that pertained to a question about Bangkok. Subjects were told to leave the envelope pertaining to the Bangkok question in the manila envelope and to put the envelope that had information pertaining to the Boston question onto the desk in front of them. Half of the subjects were endowed with envelopes that pertained to question 1 and the other half were

[2] The full set of experimental instructions are available from the author upon request.

Part 2

- Median answer for the 116 Massachusetts Institute of Technology students who answered the question:

1. From 1961 through 1990, what was the average high daily temperature in January in Boston?

Figure 2-1: Envelope containing information pertaining to question 1

endowed with envelopes that pertained to question 2. Three sessions of 14 subjects each were endowed with information about temperature and the other three sessions of 14 subjects each were endowed with information about gas prices.

Each envelope contained the median answer given by 116 MIT students who answered either question 1 or question 2. Subjects were told about the university where this data was collected as well as the date of collection. For example, the subjects who were endowed with information about question 1 had in front of them an envelope like the one in Figure 1.

The information in the envelopes was collected as part of a separate study (Healy (2004)). In that study, I asked each of the four questions to 116 introductory economics students at MIT.

The first test of an information endowment effect involves giving the subjects who are endowed with information about Boston temperature to trade their endowments for information about gas prices in Bangkok. If there is no information endowment effect, then half the subjects are expected to trade their endowments. Previous research (Knetsch (1989)) has found that 89% of subjects choose to keep a coffee mug and 90% of subjects keep a chocolate bar when given the opportunity to trade, evidence of a strong consumption endowment effect. If there is an information endowment effect, less than half of the subjects are expected to trade their envelopes. If this is the case, subjects attach a higher value to information with which they are endowed than to information with which they are not endowed.

To make the test of an endowment effect valid, the pieces of information that can be traded need to be of similar value. To see if this was the case, I asked 46 people whether they would prefer to observe the median answer about Boston temperature or to observe the median answer about gas prices in Boston. Of those 46 respondents, 21 preferred to observe information about Boston temperature, so that the hypothesis that half the people would choose information about temperature and the other half would choose information about gas prices cannot be rejected. This test showed that the two pieces of information were sufficiently close in value that they could be used in a trading exercise.

After I explained to subjects the types of available information, subjects who were endowed with information about question 1 (question 2) were given the opportunity to trade their endowment for information about question 2 (question 1). I and an assistant approached each carrel and placed an envelope pertaining to the other question next to the endowment envelope. Each subject indicated which piece of information they wished to keep by pointing to her chosen envelope. The unchosen envelope was taken away from each carrel.

Once trading was complete, subjects opened their envelopes and provided final answers to each of the questions. The ways in which subjects updated their answers is discussed in Section 4.

2.2.3 Part I Results

The evidence from the trading treatments provides strong evidence against an information endowment effect (Table 1). Of the 84 subjects who participated in the experiment, 43 (51%) traded their endowment envelope for an envelope that contained information about the other question. Overall, there was a small preference for information about temperature. 24 (57%) of the subjects who were endowed with information about temperature traded their endowments for information about gas prices. Only 19 (45%) of the subjects who were endowed with information about gas prices traded their endowments for information about temperature.

These results indicate that, in contrast to the previously cited research that documents the presence of endowment effects when subjects are endowed with goods like mugs or chocolate bars, subjects do not show an information endowment effect. This is evidence consistent with the idea that the endowment effect does not cause subjects to open a separate mental account when they are endowed with information.

One interesting element of heterogeneity emerges from the trading data that is also repeated in the auction treatments in the next section. The propensity to trade is significantly higher for men than for women. Table 2 describes the discrepancy for men and women. While 19 of the 29 men (66%) traded, only 24 of the 55 women (44%) traded their envelopes. In aggregate, the larger share of women in the sample led to the overall trading volume being closer to 50%.

The experimental design that was implemented here was meant to eliminate any possible spurious causes of an endowment effect. In contrast to to other endowment effect experiments, any minimal transaction cost was eliminated by having subjects point at the envelope they wished to choose. Placing the second envelope in the carrel and asking the subjects to point at their preferred envelope eliminates non-trading as a default option. As discussed earlier, the subject possesses the endowment envelope from the start of the experiment, giving the subject significantly longer to form an attachment to that envelope as opposed to the one for which she could trade.

To provide a second test for the presence of information endowment effects, I also conducted an auction exercise for information relating to the two questions about Bangkok. The results that I find there are broadly similar to what I find with the trading treatments. Interestingly, the pattern of a greater preference for endowed information among women as compared to men also carries over to the auction treatment.

2.3 Experimental Design II

2.3.1 Information auctions

After the trading experiment was completed, the same subjects participated in a treatment which elicited the values that they attached to information with which they were endowed and to other information. The subjects were told to take the other small envelope out of the manila envelope and put it on their desk in front of them. They were told the envelope that they had been given contained the median answer given by 130 students at Thammasat University (a top-ranked university in Bangkok) who answered questions 3 and 4. Subjects were told about the university where this data was collected as well as the date of collection. Everyone in a given session was endowed with the same information. Half of the sessions and half of the participants involved subjects who were endowed with question 3 information; the other half involved subjects who were endowed with question 4 information.

Again, I conducted a test to check to see if information pertaining to question 3 was of similar value to information pertaining to question 4. I asked the same 46 people who stated their preferences for question 1 versus question 2 whether they would prefer to observe the median answer about Bangkok temperature or to observe the median answer about gas prices in Bangkok. Of those 46 respondents, 24 preferred to observe information about Bangkok temperature, so that the hypothesis that half the people would choose information about temperature and the other half would choose information about gas prices cannot be rejected. This test showed that the two pieces of information were sufficiently close in value that they could be used in the auction treatments.

Subjects were again told both verbally and on their envelopes that the piece of information available to them on a given question was the median answer given by the Thai students. The subjects were also informed that the Thai students were asked to answer the questions in units that were familiar to them and that the median answer had been translated into units that were familiar to Loyola Marymount students.

Before observing any information pertaining to either question 3 or question 4, subjects were informed that they were then going to participate in two auctions: one auction in which they could buy information and another in which they could sell information. Subjects who were endowed with information about question 3 (question 4) were told that they would be able to bid on the opportunity to buy the median Thammasat answer for question 4 (question 3) and that they would also participate in an auction in which they could sell the information that they had been given that pertained to question 3 (question 4).

In the buying auction, subjects would provide the price that they would be willing to pay (WTP) for information about question 4 (question 3). All of the bids would be collected and the five highest bidders would be declared the winners of the auction. They would receive envelopes with information pertaining to question 4 (question 3) and would pay the amount of the sixth highest bid. Subjects were given a \$5 budget for buying information and told that any unused amount from this budget would be added to their payments. In the selling auction, subjects would provide the price that they would be willing to accept (WTA) to sell the information they had been given about question 3 (question 4). All of the selling bids would be collected and the five lowest bidders would be declared the winners of the auction. They would hand in their envelopes and each would receive the amount of the sixth lowest bid.

Similar to List's (2003) field experiments with sports memorabilia, the subjects were told that, under this mechanism, it was in their interest to choose selling and buying prices that reflected the actual values that they put on the information. This fact was demonstrated through a simple practice auction that demonstrated how the auction mechanism worked.

When practice was complete, each subject wrote down her WTP for question 4 (question 3) information and her WTA for her endowed question 3 (question 4) information. Each subject did this by checking off her selected bid from the choices that ranged from \$0.25 to \$5.00 in increments of \$0.25. Then the bids were collected and the winners were announced, first for the buying auction and then for the selling auction. Some subjects ended up with no information, some had information about only one question, and others had information about both question 3 and question 4. After all transactions were complete, subjects were instructed that they could open their envelopes and examine the information. Subjects then gave their final answers to questions 3 and 4.

2.3.2 Part II results

The buying and selling prices that the subjects chose provide further evidence against the presence of a significant endowment effect. In keeping with the results from the trading experiment, the female subjects put higher relative value on endowed information than do the male subjects. For women, in fact, the selling price for information is significantly higher than the buying price, although the difference in prices is not large. In the overall data, there is no significant difference between the average buying and selling prices. The average chosen buying bid is slightly lower than the average chosen selling bid.

Table 3 describes the general results from the auctions. The average buying price (WTP) was \$2.49 and the average selling price (WTA) was \$2.81. The difference between the two is not significant at conventional levels. These results contrast with the large and significant differences between WTP and WTA that have been found when subjects are endowed with goods.

The general equality between the mean WTP and WTA in the overall sample extends to the two treatments. For the group that was endowed with the median Thai answer about the Bangkok temperature (question 3), the mean WTP was \$2.60 and the mean WTA was \$2.88. For the group endowed with gas price (question 4) information, the mean WTP was \$2.38 and the mean WTA was \$2.74. In each case, the difference between the WTP and WTA was not significant, so that the null hypothesis of no endowment effect cannot be rejected.

The small difference between the mean WTA and the mean WTP comes entirely from the female subjects in the sample. Table 4 shows the breakdown of WTP and WTA values by gender. Male subjects actually had a slightly larger WTP than WTA, although the difference is insignificant (\$2.68 compared to \$2.51). Female subjects had a significantly smaller mean WTP compared to mean WTA (\$2.38 compared to \$2.97). This corresponds to the trends that were observed in the trading treatments. Relative to the male subjects, the female subjects put higher value on endowed information compared to the other available information. In the aggregate, in each case, the null hypothesis of no information endowment effect cannot be rejected. In the auction treatment, the WTA is significantly larger for women, although the difference between the WTA and the WTP is still much smaller than the differences found when subjects are endowed with goods such as mugs.[3]

While the average WTP and WTA values are generally of similar magnitude, some interesting behavior underlies this similarity. Specifically, a large group chose \$3 as the selling price. Of the 82 subjects who participated in the auctions, 18 (22%) chose exactly \$3. In both the buying or selling auctions, no other price was chosen half as frequently. Figures 2 and 3 show the full distribution of chosen buying and selling prices. Subjects appear to perhaps think more carefully about their buying prices then they do about the selling prices. The tendency to settle on a whole dollar amount may reflect a coarser valuation of the endowed information.

[3] As a rough rule, endowment increases valuation by a factor of two in the cases of goods such as mugs (Knetsch (1989), Kahneman, Knetsch, and Thaler (1990)).

It is still clear that, while it seems subjects treat selling somewhat differently than they do buying, there is no systematic difference between WTP and WTA values. Endowment does not increase how the subjects value information.

2.4 Discussion

2.4.1 How information was used

This experimental design, in contrast to previous work that looked at goods, makes it possible to investigate how endowment affects utilization. Does being endowed with information cause an individual to weight it differently? Other research has found that agents are more willing to not view a movie if it they were given it for free than if they had paid for it ( ). Agents were less willing to ignore sunk costs when they paid for the item that they would be returning without using. This research makes it possible to specifically answer the question of how endowment affects utilization. One way to get at this question would involve seeing if an agent uses a coffee mug more when she is endowed with it rather than paying for it.

Here, it is possible to test the hypothesis that subjects use information the same whether they were endowed with it or they had to pay for it. This involves making the assumption that subjects' buying and selling prices reflect their valuations in the same way. The close correspondence between the mean WTP and WTA values justifies this assumption. Given this assumption, we can look at how the people who choose high valuations use the information for a given question. For example, consider a subject who values a question's information at \$4.00. If the subject is endowed with the information, her selling bid would fail. If she is not endowed, her buying bid would be successful and she would acquire the information. We can compare how the endowed person treats the information compared to the unendowed person.

Table 6 describes the median answers that were available to be seen. For each question, the median answer that a subject could have observed is quite close to the correct answer. A significant number of subjects who observed information simply chose the observed median when picking a final answer, both for question 3 and for question 4. Of those 49 subjects who kept their endowed information by failing to win the selling auction, 25 (51%) chose the answer they observed. Of those 32 subjects who obtained information by winning the buying auction, 15 (47%) chose exactly the answer that they observed.

More generally, the way in which subjects update their answers upon observing information can be captured by defining the magnitude of the update to be:

$$Weight\ given\ to\ observed\ median = \max\left(1, 1 - \left|\frac{Final\ answer\ -\ Observed\ median}{Initial\ answer\ -\ Observed\ median}\right|\right)$$

The weight is bounded from above at one to account for subjects like the one who initially answered 74°F for question 3, observed the Thai median of 90°F and then gave a final answer of 93°F. The weight captures the extent to which subjects change their answers after observing information. If a subject does not change her answer, the estimated weight is zero. If a subject chooses the observed median, the estimated weight is one.

Table 7 shows the average weights for subjects who observed their endowed information since their selling bids failed and for subjects who observed information that they purchased. The table shows that the information is used in similar ways whether it is endowed or purchased. The average weight that endowed subjects use is 0.87, compared to an average weight of 0.91 for subjects who purchased their information. Due to the fact that there were only five winners (out of 14 subjects) in the buying and selling auctions in each session, the average winning buying price is somewhat higher than the average winning selling price. To ensure that information is treated in the same way independent of whether it is endowed or purchased, the second row in Table 7 looks at how information is treated for the 22 highest buying bids and 22 highest selling bids. In each case, subjects put an average weight of 0.89 on the observed information.

The results make it clear that endowment does not affect how the information is used. Subjects use the information they observe in nearly identical ways whether the information is purchased or endowed.

2.4.2 Subject earnings and information valuation

The experimental design also makes it possible to look at whether or not subjects appropriately value the information available to them. The results here are only suggestive, since the information available to subjects happens to be quite close to the correct answers, as is seen in Table 6. This should not be surprising to the subjects, since it could be reasonably expected that residents of Boston or Bangkok would have precise knowledge of gas prices and temperature there.

Subjects whose initial answers to the questions are further from the correct answers value the available information somewhat higher than do the subjects with more accurate initial answers, however the differences are not significant. Table 8 displays the results of regressing the distance between subjects' initial answers and the correct answer on the revealed valuation of that information. For each question, there is no significant difference. The point estimates for question 3 predict that a subject whose answer is 2 degrees further from the correct answer of 90°F would value the information at \$1 higher. The point estimates for question 4 indicate that a subject whose answer is \$0.10 further from the correct answer of \$1.42 would place \$1 higher value on information pertaining to that question.

The results of the auction confirm that the auctions, in particular the buying auctions, work well. Table 9 shows the results from regressing subjects' total payments on their buying bids, their selling bids, and both of these. In each case, the estimated coefficients on the buying bid and selling bid variable are close to zero and the hypothesis that information valuation is independent of subjects' payments cannot be rejected. The coefficient is almost exactly zero in the case of the buying auctions, corresponding to the earlier empirical finding that subjects were more willing to carefully specify their buying prices than their selling prices. In both auctions, the evidence shows that the markets find an efficient price. The buying and selling prices are set such that, whether or not they win the auction, subjects on average do equally well.

This finding that the winning bids are efficient confirm the earlier findings that there is no information endowment effect. When subjects are endowed with information, markets work well and behavioral anomalies do not occur.

2.5 Conclusion

This paper reports the results of an unusual experiment that endowed subjects with information rather than goods. Previous experiments have provided evidence suggesting the existence of an endowment effect for goods. If correct, this anomaly calls into question the assumptions underlying our most basic models of consumer behavior. Given the far reaching implications of the endowment effect, it is important to understand what mechanism underlies this behavioral anomaly.

The results in this paper point towards the endowment effect operating on preferences as opposed to judgment. If an agent values a good such as a coffee mug at twice as high a price if the mug is part of her endowment, either she could believe that the coffee mug is actually worth twice as much to the average consumer or she could understand that the mug was only more valuable to her without having her perceptions of its general value being affected. In the latter case, the agent's preferences are affected by endowment; in the former case, the agent's judgment is affected.

The experimental design eliminated the possibility of the endowment effect operating on preferences. Since the information could not be consumed, it does not enter the utility function. Any information endowment effect could only be reflected through judgment. Someone could irrationally believe that their information was better than it was because it belonged to them. The results reported here show that this is not the case. No information endowment effect arises. Here, at least, endowment does not distort judgment. This provides evidence pointing towards the endowment effect exerting influence only on agents' preferences and not on their judgment.

Bibliography

[1] Genesove, David and Christopher Mayer (2001). "Loss Aversion and Seller Behavior: Evidence from the Housing Market," *Quarterly Journal of Economics*, CXVI, 1233-1260.

[2] Healy, Andrew (2004). "How Does an Advisor's Nationality Affect the Weight Given to Her Advice? Experimental Evidence from Thailand," MIT working paper.

[3] Kahneman, Daniel and Amos Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, XLVII, 263-291.

[4] Kahneman, Daniel, Jack Knetsch, and Richard Thaler (1986). "Fairness as a Constraint on Profit-Seeking: Entitlements in the Market," *American Economic Review*, LXXVI, 728-741.

[5] Kahneman, Daniel, Jack Knetsch, and Richard Thaler (1990). "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, XCVIII, 1325-1348.

[6] Knetsch, Jack (1989). "The Endowment Effect and Evidence of Nonreversible Indifference Curves," *American Economic Review*, LXXIX, 1277-1284.

[7] List, John (2003). "Does Market Experience Eliminate Market Anomalies?" *Quarterly Journal of Economics*, CXVIII, 41-72.

[8] Odean, Terrence (1998). "Are Investors Reluctant to Realize Their Losses?" *Journal of Finance*, LIII, 1775-1798.

[9] Shea, John (1995). "Union Contracts and the Life Cycle - Permanent Income Hypothesis," *American Economic Review*, LXXXV, 186-200.

[10] Thaler, Richard (1980). "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior and Organization*, I, 39-60.

[11] Tversky, Amos and Daniel Kahneman (1991). "Loss Aversion in Riskless Choice: A Reference-Dependent Model," *Quarterly Journal of Economics,* CVI, 1039-1061.

Table I
Summary Statistics for the Trading Experiment

Variable	Percent traded	p-value for test of trading volume = 50%
All subjects (n=84)	51%	
Endowed with question 1	57%	0.83
Endowed with question 2	45%	

a. Question 1 refers to the January temperature in Boston.
b. Question 2 refers to gas prices in Boston.

Table II
Summary Statistics for the Trading Experiment, by Gender

Variable	Percent traded	p-value for test of trading volume = 50%
All subjects (n=84)	51%	0.83
Male subjects (n=29)	66%	0.1
Female subjects (n=55)	44%	0.35

Table III
Summary Statistics for Auction Treatments

A. Mean buying and selling prices

Variable	WTP mean	WTA mean	*p*-value for test of mean WTP = mean WTA
All subjects (n=82)	\$2.49	\$2.81	0.12
	(1.22)	(1.29)	
Endowed with question 1 (n=40)	\$2.60	\$2.88	0.33
	(1.16)	(1.28)	
Endowed with question 2 (n=42)	\$2.38	\$2.74	0.23
	(1.27)	(1.31)	

B. Median buying and selling prices

Variable	WTP median	WTA median	*p*-value for test of equality of distributions
All subjects (n=82)	\$2.50	\$3.00	0.09
Endowed with question 1 (n=40)	\$2.75	\$3.00	0.49
Endowed with question 2 (n=42)	\$2.13	\$3.00	0.1

a. Question 3 refers to the January temperature in Bangkok and Question 4 refers to gas prices in Bangkok.
b. Standard deviations are in parentheses.
c. Two subjects in the first session refused to participate in the auctions.

Table IV
Summary Statistics for Auction Treatments, by Gender

Variable	WTP mean	WTA mean	*p*-value for test of mean WTP = mean WTA
All subjects (n=82)	\$2.49	\$2.81	0.12
	(1.22)	(1.29)	
Male subjects (n=29)	\$2.69	\$2.52	0.63
	(1.18)	(1.27)	
Female subjects (n=53)	\$2.38	\$2.97	0.02
	(1.23)	(1.28)	

a. Standard deviations are in parentheses.

Table V
Winning Selling and Buying Prices, by Session

Session	Winning buying price	Winning selling price
1	\$3.00	\$2.50
2	\$2.00	\$2.25
3	\$2.75	\$3.00
4	\$3.50	\$3.00
5	\$3.50	\$2.50
6	\$2.50	\$2.25

a. In sessions 1, 3, and 5, subjects were endowed with information about Bangkok temperature.
In sessions 2, 4, and 6, subjects were endowed with information about gas prices in Bangkok.
b. The winning buying price was the sixth-highest bid; the winning selling price was the sixth-lowest bid.

Table VI
Median and Correct Answers to the Questions

Question	Median answer given by MIT or Thammasat students	Correct answer
A. Questions about Boston		
1	35°F	36°F
2	\$1.80	\$1.90
B. Questions about Bangkok		
3	90°F	90°F
4	\$1.51	\$1.42

a. Questions 1 and 3 refer to the high daily January temperature; questions 2 and 4 refer to gas prices.
b. For questions 1 and 2, the median refers to the distribution of MIT student answers.
For questions 3 and 4, the median refers to the distribution of Thammasat student answers.

Table VII
Weights Given to Purchased and Endowed Information

Session	Average weight
A. Endowed subjects	
All subjects whose selling bids fail (n = 49)	0.87
22 highest selling bids	0.89
B. Subjects who purchased information	
All subjects whose buying bids win (n = 32)	0.91
22 highest buying bids	0.89

a. In sessions 1, 3, and 5, subjects were endowed with information about Bangkok temperature. In sessions 2, 4, and 6, subjects were endowed with information about gas prices in Bangkok.
b. The winning buying (selling) price was the sixth-highest (sixth-lowest) bid.

Table VIII
Relationship Between Accuracy of Initial Answer and Information Valuation

Question	Regression coefficient
A. Question 3: Bangkok temperature	
Both buying and selling bids (n = 82)	2.04
	(1.28)
Just buying bids (n = 40)	0.91
	(1.80)
Just selling bids (n = 42)	3.05
	(1.92)
A. Question 4: Bangkok gas prices	
Both buying and selling bids (n = 82)	0.10
	(0.13)
Just buying bids (n = 42)	-0.15
	(0.18)
Just selling bids (n = 40)	0.29
	(0.20)

a. Standard errors are in parentheses.
b. For questions 1 and 2, the median refers to the distribution of MIT student answers.
For questions 3 and 4, the median refers to the distribution of Thammasat student answers.

Table IX
Total Payments and Information Valuation

Dependent variable: Total payments received by the subject

Variable	Regression coefficient (1)	(2)	(3)
Selling bid	0.27		0.27
	(0.32)		(0.33)
Buying bid		0.00	0.03
		(0.34)	(0.35)
Constant	15.27	16.04	15.19
	(1.00)	(0.96)	(1.39)

a. Standard errors are in parentheses.

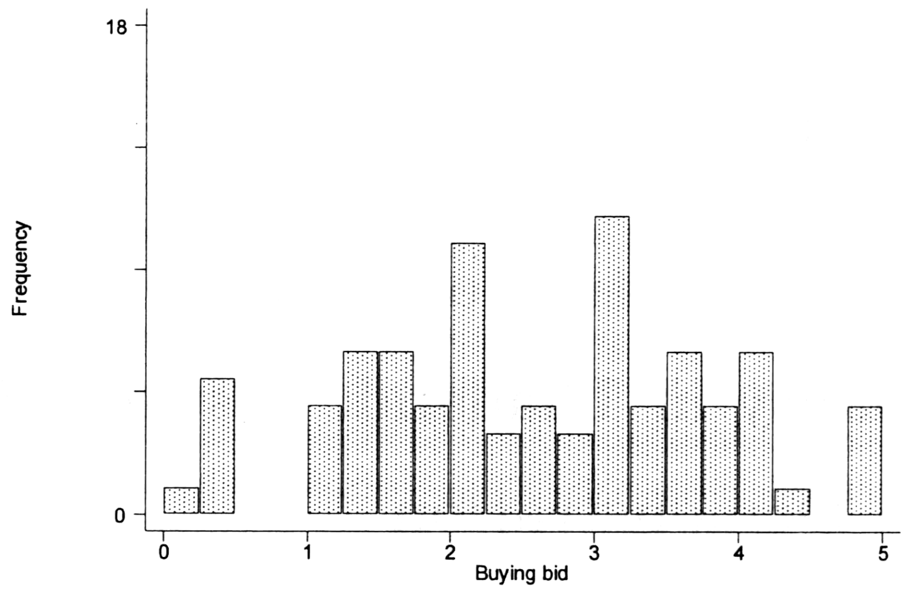

Figure 2: Histogram for chosen buying prices

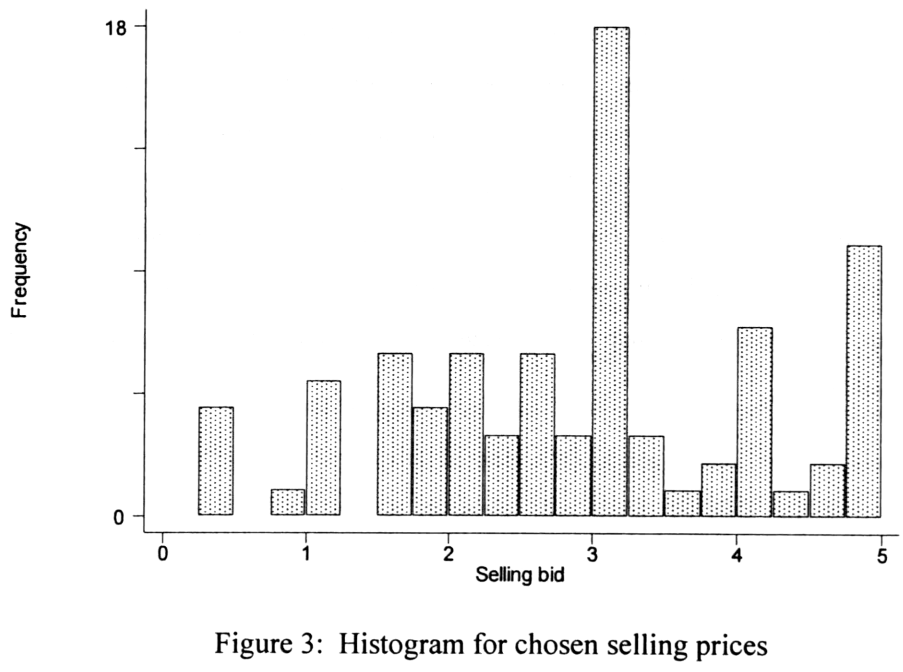

Figure 3: Histogram for chosen selling prices

Chapter 3

A Spatial Picture of Poverty and Inequality in Thailand: Estimates and a Demonstration

Summary 3 *We provide the first comprehensive estimates of Thai poverty and inequality at levels below the changwat (province). We do this by combining a household survey, which has income and consumption data, with the 2000 Census, which is representative at the level of the individual. To improve the precision of our estimates, we propose improvements to the methods previously applied in other settings. We produce estimates that correspond closely to those produced by the survey alone at the province level. In addition to passing this test, the standard errors on the district-level estimates of poverty and inequality are actually smaller than the changwat-level estimates that come from the household survey. We conclude by demonstrating the potential for this spatial decomposition to improve policy. If policymakers want to reduce the poverty gap, our results can help direct resources in a much more effective way than was previously possible.*

3.1 Introduction

Until recently, in most developing countries, available national data could only tell policymakers if an entire province is poor. Household surveys, the best available data source about household income and consumption, are generally only representative at the province level. In Thailand, for example, the changwat (province) is the stratum in the Socio-Economic Survey (SES) and the lowest attainable geographic level of disaggregation for SES-based poverty and inequality measures. There are seventy-six changwat in Thailand, a country of sixty million people. Policymakers interested in doing a better job of targeting the poor in Thailand have pointed to the need for a more precise spatial description of poverty and inequality.

This is of even greater interest due to the recent decentralization currently taking place in Thailand. Previously, the federal government directly disbursed nearly all funds, with there being little local discretion. In the last few years, Thailand has moved towards more local control of policy implementation. As a greater balance develops between local and federal authority, the ability to implement policies on a local level will increasingly be present. This makes a disaggregated description of poverty and inequality especially valuable.

In this paper, we present results that provide this information. We do this by combining the 2000 Socio-Economic Survey (SES) with the 2000 Population and Housing Census. With some important improvements, we apply the poverty mapping methodology developed in Elbers, Lanjouw, and Lanjouw (2003) to the Thai case.[1] The basic idea that we apply here is quite simple.

The strength of the household survey is the comprehensive data it provides for each household in the sample. Specifically, it reports household income and consumption data. It is, however, representative only at the level of the changwat (province). On the other hand, the Census covers all households, but it lacks data on household income and consumption.

By combining the two data sources, we can take advantage of the strengths that each possesses. We use the SES to model income and consumption as functions of correlated variables. We choose right-hand side variables that both the SES and Census report. The

[1] For papers that have combined survey and Census data to spatially decompose poverty and inequality estimates, see Hentschel et al. (2000) for Ecuador, Alderman (2002) et al. for South Africa, and Mistiaen et al. (2002) for Madagascar.

model that best fits for the SES households can then be used to predict income and consumption for the Census households. Using careful econometrics to get the correct standard errors, we obtain estimates of poverty and inequality at all levels of geographic aggregation, from the changwat all the way down to the village. The standard errors become larger for lower levels of spatial aggregation. We find that we obtain estimates with reasonable standard errors down to the tambon (subdistrict) level. The standard errors for our estimates at the *district* level are actually somewhat smaller than the estimates obtained using the SES alone at the *province* level.

Of course, a precise picture matters only to the extent that we can verify that picture's accuracy. Fortunately, we can compare our results at the changwat level against the household survey's estimates. We show that our estimates of poverty and inequality correspond closely to the changwat level estimates that the SES gives. The estimates found in the poverty map are within the 95% confidence intervals for the changwat poverty and inequality estimates from the SES approximately 95% of the time. When we look at the level of the region, of which there are five in Thailand, the map and the SES show an extremely close correspondence.

These results indicate that the map can be useful in targeting policies to geographic levels below the changwat. At levels well below 5,000 households, we produce estimates of poverty and inequality that have standard errors of the same size as those obtained at the changwat level using the SES. We now can say a tambon (subdistrict) is poor with the same confidence we could previously apply to an entire changwat.

The primary usefulness of the estimates we report is as an input to guide policy and further research. To demonstrate the value of our results, we ask a policy question. If the goal is reducing the poverty gap, how much greater of a reduction can be achieved by using our estimates as opposed to the data previously available? Given the precision with which we can estimate poverty at even the subdistrict level, the answer is not surprising. Our results have the ability to lead to a much more effective allocation of resources.

Section 2 describes the Thai data and the methodology that we apply to it. Section 3 compares our results to those produced by the SES for high levels of geographic aggregation. Section 4 reports our estimates of poverty and inequality at the amphoe and tambon level. Section 5 considers a modified model that can be used in addition to our primary estimates.

In Section 6, we demonstrate that these estimates make possible significant improvements in resource allocation. Section 7 concludes.

3.2 Poverty Mapping Methodology and the Data

3.2.1 Methodology

Our goal is to predict income and consumption for the Census households. To do this, we find variables that are present both in the Census and the household survey (SES). These variables are then used to estimate models of income and consumption for the survey households.[2] The coefficients (and their standard errors) and the resulting residuals then enable us to estimate income and consumption for the Census households.

Finding variables present in both the Census and the household survey

The first step is to identify the variables that are present in the SES and the Census, and identically defined in each data set. These variables will form the set from which we choose the variables in our income and consumption models. Before estimating those models, we run tests to confirm that variables are identically defined in the two datasets. We do this by comparing the changwat means for each variable between the Census and the SES. If a variable really is defined in the same way, then we should not reject a test of equality between the Census mean and the SES mean.

We consider household asset, demographic, and occupational variables that are plausibly correlated with income or consumption.[3] To ensure that variables are defined the same way in the Census and the SES, we check to see if the Census mean (taken to be the truth) falls in the 95% confidence interval for that variable's mean that we find in the SES.[4] We find the

[2] For India, Deaton and Taorozzi (2000) and Taorozzi (2002) use related methods to estimate consumption for the National Sample Survey (NSS) households. They have a set of variables that they know are well-measured and use them to estimate consumption for the households, leading to estimates of Indian poverty and inequality. Their goal is to resolve problems with the NSS, while ours is to obtain a spatial picture of poverty and inequality.

[3] See Appendix A for a list of all the variables that are considered in the zero stage.

[4] To be more precise, we should use more than a variable's first moment. Most of our variables will be dummies, however, making the mean a sufficient statistic for the distribution. We have performed sensitivity checks for other variables, and find that looking at higher moments adds no change to the set of variables that pass the zero stage.

variables that pass this test for each changwat, urban and rural. To avoid the possibility that any variables essentially act as a dummy for a given household or two, we only consider dummy variables that have a mean greater than 0.03 and less than 0.97. Some of the strata in the SES have less than one hundred households. For such a stratum, a dummy variable that has mean 0.99 is zero for all households, save one. This variable then acts as a dummy for that household in any model. Since the goal is to estimate income and consumption for the Census households, this is clearly not a variable that can be helpful in constructing an appropriate model. The household represented by the dummy variable cannot be identified in the Census. Even if it could, we would not want to use a degree of freedom representing one household, when we need to predict income and consumption for all the Census households.

Selection of household-level correlates

After ascertaining which variables are defined in the same way in the Census and SES, we model household income or consumption in the SES as a function of the variables that pass the zero stage. Where y_{ch} refers to log income of household h in cluster c (the cluster is the village in the SES), X_{ch} denotes the set of available regressors, and η_c are village dummies, our regression equation is

$$y_{ch} = E(y_{ch}|X_{ch}^T, \eta_c) + u_{ch} = X_{ch}^T\beta + \eta_c\gamma + u_{ch} \ . \tag{3.1}$$

We improve the previous methodology by including village dummies in the regression. In this way, we pick the household variables that best correlate with the household-level variation in income. Since later steps will account for location-level variation, our aim is to choose the variables that specifically capture household-level variation.

It is important to note that regression equation (3.1) does not describe a causal relationship. The usual concerns about endogeneity do not apply to our models. We are interested in obtaining consistent estimates of income and consumption for the Census households; we are not concerned with any of the individual regression coefficients. We assume that u_{ch} is distributed with mean zero and variance-covariance matrix Σ. A stepwise regression selects all variables that are significant at the 5% level.

There is a question as to whether the regressions should be estimated using the household weights provided in the SES. Note that the usual arguments against weighting do not apply in our context. Our regression is descriptive and thus weights may be needed to provide a consistent estimate of the population regression function (Deaton (1998)). Since a Hausman test rejects parameter equality at a 5% level for most of the provinces, we use the household weights provided in the SES when estimating our models. Table 1 summarizes the results of the tests that we conducted for consumption in the rural areas.

Selection of location-level correlates

After choosing household-level variables, we then select variables to absorb the location component of the variance. We choose from Census means at the village level for rural areas and Census means at the tambon level for urban areas.[5] For rural areas, the 1999 Village Survey provides additional candidate variables to model the effect of location on income and consumption. This is particularly helpful since poverty in Thailand, as in most developing economies, is concentrated in rural areas. We discuss this in Section 3.

We cannot use location dummies to predict income in the Census because we can only estimate parameters for the SES villages. We can certainly include village dummies in the model that we use in the SES. However, this would not be useful since we need to model income and consumption for all the Census households and only a very small number of the Census villages are contained within the SES sample.

Consequently, we approximate the location effect with location-level regressors that are available in the Census. To do this, we estimate a village effect from regression (3.1) and regress it on the set of location variables. Our regression equation, where $\widehat{\eta}_c$ is the estimated village effect from (3.1), is

$$\widehat{\eta}_c = Z_i^T \boldsymbol{\alpha} + v_{ch}, \tag{3.2}$$

where Z_i^T are the location variables, $i = t$ (tambon) for urban areas, and $i = c$ (village) for rural areas. Here we employ a forward stepwise regression procedure that ensures all the previously selected household variables remain and then adds any location variables that are significant at

[5] We cannot use village-level means due to a limitation in our ability to match village identifiers between the SES and the Census for urban areas.

the 5% level.

Modeling the residual distribution

The stepwise regressions (3.1) and (3.2) select the set of household and location variables that we use to model income and consumption. To simulate income and consumption for the Census households, we also need to consider the distribution of the residuals from the regression of log income and consumption on the selected household and location variables.

Where the superscript S indicates the set of selected variables, $X_{ch}^{T,S}$ denotes the set of selected household variables, and $Z_i^{T,S}$ denotes the set of selected location variables, we have

$$y_{ch} = X_{ch}^{T,S}\boldsymbol{\beta} + Z_i^{T,S}\boldsymbol{\alpha} + u_{ch} \tag{3.3}$$

We model the residual as having a cluster component and a household component.

$$u_{ch} = \eta_c + e_{ch}$$

Both due to the frequent presence of a significant location component and the fact that the variance of e_{ch} depends on observables, we can easily reject homoskedasticity of u_{ch}. To better simulate household income and consumption, we model this heteroskedasticity so that our simulations draw appropriately distributed error terms.

To do this, we take the residual from the regression of y_{ch} on the selected household and location variables, $\widehat{u_{ch}}$, and break it up into a cluster component and a household component. We then use a logistic model to find the conditional variance of the household component. Where e_{ch} is the household component of the residual and H_{ch}^T denotes household and location-level variables including squares, cubes, and interactions, the regression we consider is

$$\ln\left[\frac{e_{ch}^2}{A - e_{ch}^2} + \phi\right] = H_{ch}^T\theta + r_{ch}\ . \tag{3.4}$$

The constant A is $1.05 * \max\{e_{ch}^2\}$, and bounds the predicted variance between 0 and $1.05A$. We apply $\phi = 0.001$. For Thailand, we found that failing to use 0.001 or some other small constant causes too much weight to be put on a couple of households in some cases. Even with our safeguards to ensure that no variable or combination of variables enters that essentially acts as a dummy for a given household, it is still our experience that sometimes a household will have a very small residual from (3.3). With this adjustment, we find that neither households with large residuals nor those with small residuals interfere with the choice of appropriate variables for the heteroskedasticity model. The delta method and a first-order Taylor approximation gives the following estimate for the conditional variance of e_{ch}, where $B = \exp(H_{ch}^T\widehat{\theta})$ and $\widehat{\sigma}_r^2$ is the estimated variance of r_{ch}[6]:

$$\widehat{\sigma}^2_{e,ch} \approx \left[\frac{AB - \phi A}{1 - \phi + B}\right] + \frac{1}{2}\widehat{\sigma}_r^2 \left[\frac{AB(1 - \phi - B)}{(1 - \phi + B)^3}\right] \tag{3.5}$$

We use the above variance estimator to generate a set of standardized household residuals. We then draw the location component and the household component either directly from the standardized residuals or from the most appropriate parametric distribution.[7] We consider normal and t distributions for the household error. Whether one of these parametric distributions is employed or we draw from the empirical distribution has little effect on the resulting poverty and inequality estimates.

We obtain GLS estimates of the parameters and then conduct simulations to assign error terms to households. This enables us to generate estimates and confidence intervals for the Foster-Greer-Thorbecke (FGT) indices, the Atkinson inequality measures, and the Gini coefficient at any level of geographic aggregation. We focus on the headcount index and the Gini in this paper.

3.2.2 The data

The income data we use comes from the 2000 Socio-Economic Survey. This is a stratified random sample of 24,747 Thai households. There are 76 strata, one for each changwat. Each of these strata is divided into three categories: urban, sanitary district, and rural. The Census has only two categories, rural and urban, and we employ the Census definition when classifying

[6] We thank Chris Elbers for helping with this derivation.

[7] We model a location component of the residual wherever possible. If the location variables capture the location effect particularly well, the fact that the estimate of the household variance is not the true value means that we can get a negative variance estimate for the location part of the residual. For these cases, we estimate the model without a location component to the residual.

households as either urban or rural.

Providing the greater coverage we need to obtain spatially disaggregated estimates of poverty and inequality is the 2000 Population and Housing Census. The short form of the Census completely covers the population. It has data on household composition, education, and occupations, but not assets. The 20% subsample also has information on assets. We have produced Thai maps using both the 100% sample and the 20% sample. We found that the greater precision that comes from the richer set of variables in the 20% sample dominates the loss of coverage. As a result, we use the 20% subsample to construct our estimates.[8]

We break Bangkok up into four parts due to its size. The division occurs according to the picture in Figure 1. We report results that aggregate the map estimates for the four parts.

For rural areas, we also include variables from the 1999 Village Survey. The data comes from a questionnaire sent out to the village headperson in each Thai village. Included is data on irrigation, the total number of households in a village, and farm assets. In our results, we find a somewhat better fit for rural areas compared to urban areas, and the extra information provided by the Village Survey may be the reason.

3.2.3 First-stage results

Compared to similar research previously done elsewhere, we find that the set of chosen regressors can account for a larger part of the variance in income and consumption. This likely speaks, in part, to the high quality of the Thai data. The SES is known for its reliability among the household surveys available outside the industrialized countries (Townsend and Gine (2003)).

Table 2 shows the median number of regressors and the median number of households per stratum, for each of the regions. It also describes the quality of fit of the regression model in the last two columns, which describe the median and average R^2 for the models, again by region. Notice that a relatively small number of regressors are capable of accounting for a large share of the variation in rural consumption. In Thailand as a whole, the median number of included regressors is 11 with a median of 147 households in the models. On average, 70.9% of

[8] While the 20% sample gives the most precise results down to the tambon level, this should add an extra bit of caution in considering results at the village level. Using the 20% sample introduces error in the Census means, since they are now only an estimate of the population mean. We estimate that this only becomes an issue for population units of 25 households or less. Details of these calculations are available upon request.

the variation in consumption is accounted for by the regressors that are included in the models. This means that most of the variation in our simulations will come from variation in the chosen regressors and those coefficients, as opposed to that variation introduced by the error term draws.

The qualitative nature of the results seen in Table 2 for rural consumption holds true for urban areas and for income. The set of regressors, on average, account for over 70% of the variance in consumption. This ability to precisely model income and consumption leads to the small standard errors for tambon-level poverty and inequality estimates that we report in Section 4.

3.3 Region and Changwat Results: Comparisons with the SES

Our contribution is to provide estimates of poverty and inequality at geographical levels like the amphoe and the tambon. Since our work has provided the first of these estimates for Thailand,[9] we cannot check our estimates at these levels. Here, though, we show that the estimates produced by the methodology at the region and changwat level and compare to what the SES gives. The close tracking confirms that our estimates perform well.[10]

Of course, these results are not are primary goal; our primary contribution comes from the amphoe and tambon estimates that we report in Section 4. The close correspondence between our results at the changwat level and the SES estimates, though, gives us confidence about our results at lower geographic levels. In other words, the fact that our estimates match the SES at the region and changwat levels means that we can trust our estimates of poverty and inequality at the amphoe and tambon levels. The reported results refer to normally distributed error terms, but the qualitative nature of the results is the same with other error distributions.

[9] More precisely, our work provides the first of these estimates based on the directly collected, national data sets collected by the Thai government. Another data set, based on letters sent to village leaders, gives a partial list of poor villages in Thailand. The limitations of, and concerns about, this list helped to motivate Thai officials into seeking the results reported here.

[10] We focus on the headcount and the gini in the discussion. The correspondence between the map and the SES carries over to other measures, such as the Atkinson inequality index and the poverty gap. Details are available from the authors upon request.

3.3.1 Poverty

Consider first the comparison at the region level. Table 3 shows the comparisons for the headcount index for rural areas. In all cases, the map estimate of poverty is well within the 95% confidence interval generated by combining the point estimates and standard errors for both the map and SES estimates. What is particularly encouraging is that the tracking seen in Table 3 comes not from large standard errors but from a close tracking in the point estimates.

Nationally, the estimates of poverty differ by less than a percentage point. The tracking is also very close for the North, the South, and for income in the North. The tracking is somewhat less precise for the Central and for consumption in the Northeast. We will discuss some of the results for the Central in greater detail in the section on mean income.

We also find a close correspondence between our results and the SES estimates for the headcount in urban areas. Table 4 shows this comparison. Again, the national estimates differ by less than a percentage point. In addition, the estimate for the South, the Northeast, and Bangkok are all quite close. The Central estimates are somewhat lower in the SES, but the test for equality is still not close to rejecting at a 5% level. We do reject the test for equality for consumption in the North, with an asymptotic z-statistic of 1.98. In other words, the test for equality cannot be rejected for 21 of the 22 comparisons presented in Tables 3 and 4.

At the changwat level, the correspondence between the map and the SES is also encouraging. Tables (7) and (8) in Appendix A display the results for urban and rural areas for the headcount index, mean income, and the gini. At very low poverty levels, comparisons become difficult. For example, rural Samutprakan has zero poverty in the SES. This estimate has a standard error of zero, and thus we will reject equality for any non-zero map estimate of poverty.

Consider those provinces with at least 2% poverty estimates in the SES. In rural areas, 97% of the comparisons out of 63 possible accept the hypothesis of equality between the map and the SES for income. The hypothesis cannot be rejected for 90% of the comparisons out of 59 for consumption. In urban areas, 89% of the comparisons out of 61 cannot reject for income and 95% out of 56 cannot reject for consumption. If we restrict comparisons to those changwat where both the map and the SES give poverty estimates of more than 2%, the results are much the same. There are three less available comparisons for rural consumption and 89% of the comparisons cannot reject equality.

In sum, we find a very close tracking between the map and the SES estimates of the headcount index, for income and consumption in both rural and urban areas. This holds true at the national, regional, and provincial levels. The robustness of this correspondence should inspire confidence in our estimates for the headcount at the amphoe and tambon levels.

3.3.2 Mean income

For mean income, the poverty map matches the SES with less frequency. Consider Tables 5 and 6, which contain the comparisons for rural and urban areas. In both urban and rural areas, we cannot reject equality for any of the regions, but we easily reject equality for Thailand as a whole for both income and consumption.

Moreover, some of the failure to reject is driven by large standard errors rather than close tracking between the map and the SES.[11] The correspondence is closest for consumption in the rural and urban North, income in the rural south, and for income and consumption in Bangkok. The map and the SES are particularly disparate in their estimates of mean income in the urban Central and North.

Most strikingly, our estimate of mean income is higher in all cases, revealing some sort of bias in either the SES or in our estimate. While it is not possible to make any certain conclusions, this finding is consistent with a frequent criticism of household surveys: non-response among wealthy households that is not reflected in the household weights. For example, Mistiaen and Ravallion (2003) report that non-reponse is higher among wealthy households in the US Current Population Survey and thus that the survey seriously underestimates American inequality. Deaton has documented this concern for a variety of countries, most notably India.[12] This sort of non-random non-response is consistent with the finding that our estimates for mean income are higher than those produced by the SES alone.

To see this, notice that our estimates would not be biased by the presence of non-random non-response in the SES. The model of income or consumption that we estimate for the changwat is valid for the households that are included in the sample. The coefficients and

[11] These standard errors are driven in large part by the substantial heterogeneity observed among the changwat in any region.

[12] It is believed in Thailand's National Statistical Office that the SES may suffer from this problem. The NSO is presently considering a research effort into the importance of wealthy non-response in the SES.

residuals that we use to model income or consumption for the Census households continue to be valid. Since the Census has universal coverage, our estimates are not biased by non-response in the SES.

The finding that our estimates of mean income are higher than the SES estimates carries over to the changwat results, which are shown in Tables 5 and 6 in Appendix A. For rural areas, the test for equality between the map and the SES passes 89% of the time for income and 84% of the time for consumption. For urban areas, the equality test passes only 77% of the time for income and 87% of the time for consumption. In 46 of the 47 times in which the equality test fails to pass, the map estimate of income or consumption is higher than the SES estimate.

Supporting the idea that wealthy non-response is driving the discrepancies seen in Tables 5 and 6 is the fact that the poverty map matches the household survey quite closely for measures where this is not relevant (the headcount), somewhat closely when this is of lesser importance (the gini), and much less closely for the measure where this of greatest importance (mean income). All results are consistent with the idea that the poverty map performs as one would hope and that there is non-response by the wealthy in the SES.

3.3.3 Inequality

The gini coefficient is broadly the same in the comparisons between the map and the SES. For all regions, we cannot reject the hypothesis that the gini is the same in the SES and the map. Tables 7 and 8 describe all the comparisons for the gini.

The tracking is very close in most cases, and the ranking of regions is identical in the map and SES. The only note of caution to add to this is the fact that 15 of the 18 regional comparisons give higher ginis in our estimates than in the SES. Again, this is consistent with some underrepresentation of the wealthy in the SES. This may be enough to make the inequality estimates higher in our estimates than in the SES, but not significantly so.

Confirming the close correspondence seen in Tables 7 and 8 are the changwat comparisons that can be found in Appendix A. In rural areas, the equality test cannot be rejected for 95% of the provinces for both income and consumption. The equality test passes in urban areas 95% of the time for income and 97% of the time for consumption.

As was the case for the headcount, the map produces estimates of inequality that correspond very closely to the SES estimates. These results and their robustness provide the foundation for the amphoe and tambon estimates of poverty and inequality that we report in the next section.

3.4 Thai Poverty and Inequality at the Tambon Level

Figures 2, 3, and 4 report our estimates for the tambon level, in map form. These maps show clearly how much more precise a picture of Thai poverty our maps are able to paint. There are 7411 tambon in Thailand, 80.2% of which are either entirely urban or entirely rural. On average, then, the tambon level is approximately 1/60 the size of a changwat.[13]

Despite the much lower level of geographic aggregation, the standard errors at the tambon level in the poverty map are actually lower than at the changwat level in the SES. This holds true for the headcount, mean income, and the gini, both for income and consumption. Tables 9 and 10 show the comparisons both for the mean of the standard errors over the changwat and for the median.

For all variables, the map provides more precise estimates at the *tambon* level than the SES does at the *changwat* level. This is precisely the administrative level that is presently assuming greater authority in Thailand. Given the performance of the poverty and inequality measures described in the previous section, Thai officials can use this new information to target policies at the tambon level with similar confidence to the changwat-level targeting that was previously possible.

In addition, we use Geographic Information Systems (GIS) maps to report our tambon estimates of poverty and inequality. Figure 1 shows the headcount for rural consumption, figure 2 shows the headcount for urban consumption, and figure 3 shows the gini coefficient for rural income.

Appendix B contains other GIS maps for the changwat level that are somewhat less visually taxing. These maps and the estimates behind them provide a more precise picture of poverty

[13] Specifically, there are 5498 tambon that are only rural, 447 that are only urban, and 1466 tambon that have both urban and rural parts, for a total of 8877 tambon map estimates. This compares to 151 map estimates at the changwat level, of which 75 are rural and 76 are urban.

and inequality in Thailand than was previously available. As shown in Tables 9 and 10, the complete coverage of the Census enables us to achieve disaggregated estimates of poverty and inequality with no loss in statistical precision.

3.5 Checking Our Estimates with Field Research

Despite the accuracy of our estimates at the region and province level, we attempted to check the poverty map results with field research. To do this, we conducted field visits to villages in Nakhon Sithammarat, a province in southern Thailand. Throughout this paper, we have emphasized our results at the amphoe and tambon levels. At the level of the village, not surprisingly, our estimates are considerably less precise. We feel, as a result, that our estimates are primarily useful at the amphoe and tambon levels.

Still, the observations of the field team largely accorded with the predictions of our estimates, even at the village level. The instances in which our estimates appeared not to match the actual picture also provided an opportunity for us to investigate an alternative model for more precise prediction. Specifically, we found one village where an important variable did not make it into our models.

This occurs, in part, due to the fact that we run our regression models for the changwat as a whole. In some cases, it may be possible to run our models for an amphoe instead. For some large amphoe, the SES has enough households to allow estimation of income and consumption models. In the few of the cases where the map fails to match the SES, we find that estimating a model at the lower geographic level may provide the answer to solving the problem.

3.5.1 Validation at the village level

Before the poverty map, the Thai government compiled a partial list of poor villages using a data set called the NRD2C that came from responses to surveys sent out to village leaders. We found that this list appears to perform well most of the time, but that problems such as outdated data hamper the old list. For example, the list classifies some villages as poor in part due to their lack of irrigated farmland, even though most villagers had ceased to be farmers. We also found that the list suffers from some inaccuracies in the reporting that seem to be intentional, possibly to manipulate the allocation of funds.

By comparison, the poverty map has several advantages, despite the fact that its primary value is at the amphoe or district level, where standard errors remain small. First, it accounts for all localities that might be absent from the NRD2C list because of the different or outdated categorization of rural villages and urban communities or unreturned questionnaires from village leaders. Second, it takes advantage of data that the government directly and comprehensively collects (the SES and the Census). Perhaps due to these strengths, the poverty map appears to perform at least as well as the NRD2C at measuring poverty in the validation villages, despite a lack of any direct income or consumption data at the village level.[14]

The poverty map does not, however, give the correct picture in all the villages. Since the map standard errors for mean income/consumption, the headcount, and the gini all become large for the village-level estimates, this is not surprising. Still, we attempted to ascertain why the model gave a poor prediction in some of the villages. Our preliminary investigation revealed that the problem may arise because the income/consumption models do not use some village characteristics that were relevant in predicting the general income of the villages. Most notably the models do not treat adequately well the fact that many villages in the Southern region grow para-rubber trees, and prices of raw rubber are always very important in determining the average income of the villagers. This problem could be overcome by improving the models or using other secondary data.

In the following section, we explore the former of these possibilities. The comparisons reported in this paper show that the poverty map generally performs well throughout Thailand. Finding a more refined model of income or expenditure might be capable of solving any problems that remain.

3.5.2 Using a model at a lower geographic level

Until now, we have discussed estimates of poverty and inequality that have come from a model of consumption that applied to an entire changwat. The regression coefficients that generated each household's estimated income and expenditure were the same throughout the changwat.

[14]The NRD2C asks villages leaders about the income levels in the village. When compiling the list of poor villages, this is one of the factors that the NESDB takes into account.

But, if one district in a changwat produces rubber and another produces textiles, the variables that correlate with consumption in each are likely to be different in each district. So we end up choosing between sample size and model precision in our regression models. We have already attempted to deal with this by allowing for conditional heteroskedasticity. Here, we also consider using a regression model based at the amphoe level rather than the changwat level.

The close correspondence between the poverty map and the household survey at the changwat level makes it feasible for us to look at income and expenditure models for some large amphoe, to supplement the results from the changwat models that we have already discussed. Within some provinces, enough of the households are located within one amphoe to enable us to run a regression at that level. We cannot generate changwat-level estimates to compare with the household survey. Earlier, though, we generated amphoe-level estimates using the changwat-level regressions. We can compare those estimates with what we obtain by running amphoe-level regressions.

We consider an amphoe model for one amphoe in each of four rural changwat and four urban changwat.[15] Table 11 shows the comparisons for the headcount between our original results from the changwat model and the household survey, at the level of the changwat. Table 12 shows the comparisons for the headcount between our original results and the results obtained by using an amphoe model, at the level of the amphoe.

Consider Chachoengsao. This is one province where the poverty map and the SES disagree quite sharply at the changwat level. For the amphoe that we look at in Chachoengsao, the amphoe model gives a result that would lead to closer correspondence between the map and the SES. Our original estimate of the headcount was higher than the SES, and the amphoe model leads to a lower estimate of the headcount for the amphoe than does the changwat model.

On average, the smaller sample size for the amphoe model and its greater precision cancel, and we get similar standard errors whether we use the changwat or the amphoe model. In the few places where the map struggles, consideration of more precise models offers hope that the poverty map methods can still yield accurate estimates of poverty and inequality at the

[15]The rural changwat are Amnatcharoen, Kamphaengphet, Phuket, and Samutprakan. The urban changwat are Chachoengsao, Lampang, Loei, and Satun. Here we just consider consumption poverty. Additional results for mean income and the gini coefficient are available from the authors upon request.

amphoe and tambon levels. If Thai officials consider suspect some survey data from one part of a changwat, they can still generate poverty map estimates for other amphoe by looking at more precise income and consumption models.

3.6 A Demonstration

In this section, we consider an example of how policymakers could use the estimates reported in this paper. Specifically, we ask: If the goal is reduction of the poverty gap, how much of an improvement in resource allocation does our estimates make possible? We find that our estimates make possible a much more efficient allocation of resources and that they can guide even simple policies in subtle ways.

The poverty gap is the total distance between household income and the poverty line, summed over all households with income below the poverty line for annual income or consumption. Where H is the number of households, i indexes the households, x_i is household income (or consumption), and z is the poverty line, the poverty gap in changwat c is

$$PG_c = \frac{1}{H}\sum_{i=1}^{H}(1-\frac{x_i}{z})\ 1\ (x_i > z)\ .$$

Consider a policymaker who wants to reduce the poverty gap as much as possible with \$1 of resources. We consider the marginal dollar to simplify the policymaker's problem. Assume her utility to be the expected reduction in the poverty gap. The policymaker achieves a reduction of \$1 in the poverty gap if she can target the dollar to any poor household, where we ignore the possibility of households being within \$1 of the poverty line.

Here we consider three kinds of possible targeting: changwat-level, amphoe-level, and tambon-level. Changwat-level targeting involves the \$1 being spread evenly to all the households in the changwat. Amphoe-level targeting involves the \$1 being spread evenly to all the households in the amphoe, with tambon-level targeting having the analogous definition.

Using the changwat-level targeting possible with the SES alone, a policymaker would target the changwat with the highest headcount. This clearly indicates that tambon-level targeting will be optimal in all cases in this model. Since the poorest tambon must be at least as poor as the changwat as a whole, targeting the poorest tambon will lead to at least as large an expected reduction in the poverty gap as targeting the changwat as a whole.

In reality, there are likely to be costs associated with attempting to implement policy at a lower geographic level. Corruption, most notably, may become more severe. These costs may outweigh the benefits of more precise targeting.

To model this, we consider iceberg costs of more precise targeting that are analogous to transport costs in trade theory. If the policymaker chooses amphoe-level targeting, ϕ_A of the \$1 is lost to corruption. If tambon-level targeting is chosen, $\phi_A + \phi_T$ of the \$1 is lost to corruption.

Our spatial decomposition of poverty estimates makes possible different kinds of policies for different places. Some changwat have relatively uniform poverty rates. For these changwat, the benefits of targeting at the tambon level are relatively small. Other changwat have large heterogeneity in poverty rates, and thus large benefits can be obtained by focusing resources on the poorer areas.

If corruption costs are the same across changwat, then, very different policies will be optimal in different places. Our estimates can tell policymakers which district or subdistrict is the best place (among districts or subdistricts) to which to target resources. Beyond that, our estimates can also help policymakers decide whether it is worth it to attempt targeting at lower geographic levels. To find out how our estimates can improve policy aimed at reducing the poverty gap, we compare our estimates for the headcount ratio at the changwat, amphoe, and tambon levels. The difference between the changwat-level headcount estimate and the amphoe-level headcount for the poorest amphoe gives the benefit of targeting at the amphoe level compared to the changwat level.

Likewise, the difference between the amphoe-level headcount estimate for the poorest amphoe and the tambon-level headcount for the poorest tambon gives the benefit of targeting at the amphoe level compared to the tambon level. If the benefits from more precisely targeting outweigh the corruption costs, the policymaker should target to the lower level.

Here we focus on the results for the Southern region of Thailand, which consists of 17 changwat. We run 200 simulations to estimate the distribution of the improvements in effectiveness achieved by targeting at the amphoe level compared to the changwat level, and at the tambon level compared to the amphoe level.

Table 13 shows the potential improvements for consumption and Table 14 reports the corresponding estimates for income.

In general, we find that the potential improvements are larger for poorer places, which is not surprising. A changwat with near-zero poverty will have near zero poverty in its amphoe and tambon, and so there is less to be gained by targeting at the amphoe or tambon levels. For income in Narathiwat, we would expect to achieve a further reduction in the poverty gap of 16.7% by targeting at the amphoe level rather than the changwat level. We could achieve an additional improvement of 13.6% by targeting tambon rather than amphoe.

Other results show that sometimes the primary improvement is obtained by targeting at the amphoe level and other times the primary improvement is obtained by targeting the tambon. In Yala, we would achieve three times the improvement by going from changwat to amphoe (17.0%) as we would achieve by going from amphoe to tambon (5.5%). On the other hand, in Pattani, we would achieve a much larger improvement by going all the way to the tambon level (a 17.7% improvement compared to a 7.4% improvement).

Moreover, while the greatest potential for the map to improve resource allocation lies in the poorest areas, there are other changwat where the map can lead to large improvements in policy. Take Ranong, for example. Even though the headcount estimates for the changwat is only 9.6%, we can achieve a total improvement of 27.9% by targeting policy to the tambon instead of the changwat (with 22.8% of the improvement coming from moving from the amphoe to the tambon level). In other words, $1 directed to Ranong at the changwat level would reduce the poverty gap by only 9.6 cents. Using our estimates to direct $1 to Ranong at the tambon level would reduce the poverty gap by 37.5 cents. Our estimates improve resource allocation in this case by a factor of four.

3.7 Conclusion

This paper provides the first comprehensive estimates of Thai poverty and inequality at a geographic level below the changwat (province). We do this by applying the methods developed in Elbers, Lanjouw, and Lanjouw (2003), with some improvements that have proven helpful in the Thai case. We combine household survey data with Census data to utilize the income data in the former and the complete coverage of the latter. To confirm that they are accurate, we compare our estimates of poverty and inequality at the changwat level to those provided by the household survey alone. These tests confirm that the models used to predict income and consumption for the Census households are appropriate ones.

These estimates of poverty and inequality are not an end in themselves, but instead form a base for policy targeting and future research. Since the late 1980s, Thailand has experienced a rapid rise in inequality. In the late 1990s, Thailand also endured a severe recession arising from the financial crisis. The disaggregated estimates of poverty and inequality presented in this paper offer a new source of information for identifying much more precisely those groups most affected by these events. We hope that policymakers will take full advantage of the new poverty and inequality estimates to improve policy targeting and that researchers will use them to help with future micro studies in Thailand.

Bibliography

[1] Alderman, Harold, et al. (2002), "How Low Can You Go?: Combining Census and Survey Data for Mapping Poverty in South Africa," *Journal of African Economics*, vol. 11, no. 2, 169-200.

[2] Bourguignon, Francois (1979), "Decomposable Income Inequality Measures," *Econometrica*, vol. 47, no. 7, 901-920.

[3] Deaton, Angus (1998), *The Analysis of Household Surveys*, Baltimore, MD. The Johns Hopkins University Press.

[4] Deaton, Angus, and Taorozzi, Alessandro (2000), "Prices and Poverty in India," RPDS working paper no. 196, Princeton University.

[5] Elbers, Chris, Jean O. Lanjouw, and Peter Lanjouw (2003), "Micro-level Estimation of Poverty and Inequality," *Econometrica*, vol. 71, no. 1, 355-364.

[6] Elbers, Chris, Jean O. Lanjouw, and Peter Lanjouw (2000), "Welfare in Villages and Towns: Micro-measurement of Poverty and Inequality," Tinbergen Institute Working Paper.

[7] Elbers, Chris, et al. (2004), "Unequal Inequality of Poor Communities," World Bank policy research working paper no. 3313.

[8] Hentschel, Jesko, et al. (2000), "Combining Census and Survey Data to Study Spatial Dimensions of Poverty: A Case Study of Ecuador," *World Bank Economic Review*, vol. 14, no. 1, 147-165.

[9] Mistiaen, Johan and Ravallion, Martin (2003), "Survey Compliance and the Distribution of Income," World Bank policy research working paper no. 2956.

[10] Mistiaen, Johan, et al. (2002), "Putting Welfare on the Map in Madagascar," World Bank African region working paper no. 34.

[11] Taorozzi, Alessandro (2002), "Estimating Comparable Poverty Counts from Incomparable Surveys: Measuring Poverty in India," RPDS working paper no. 213, Princeton University.

[12] Townsend, Robert and Gine, Xavier (2003), "Evaluation of Financial Liberalization: A General Equilibrium Model with Constrained Occupation Choice," World Bank Policy Research Working Paper no. 3014.

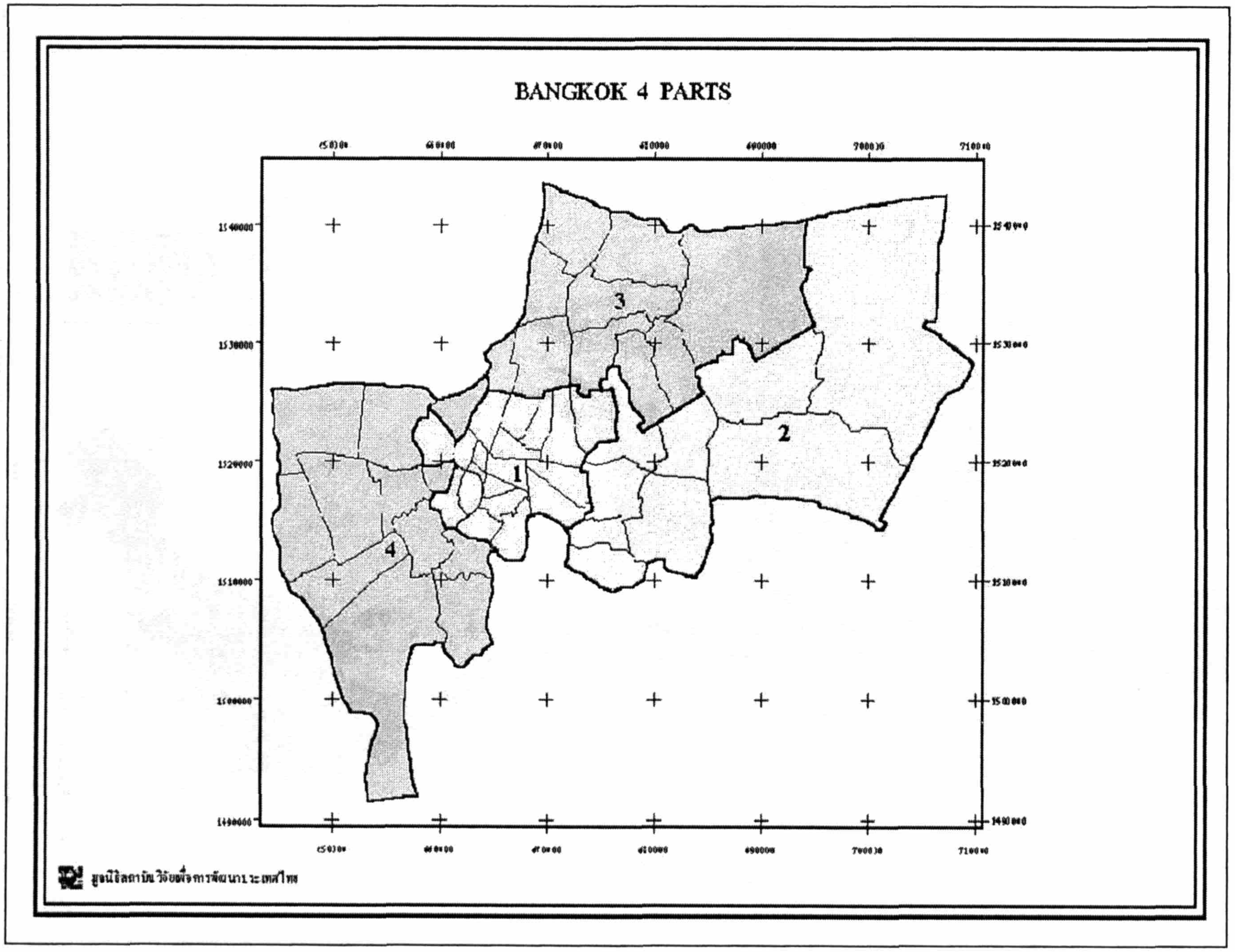

Figure 1: Decomposition of Bangkok

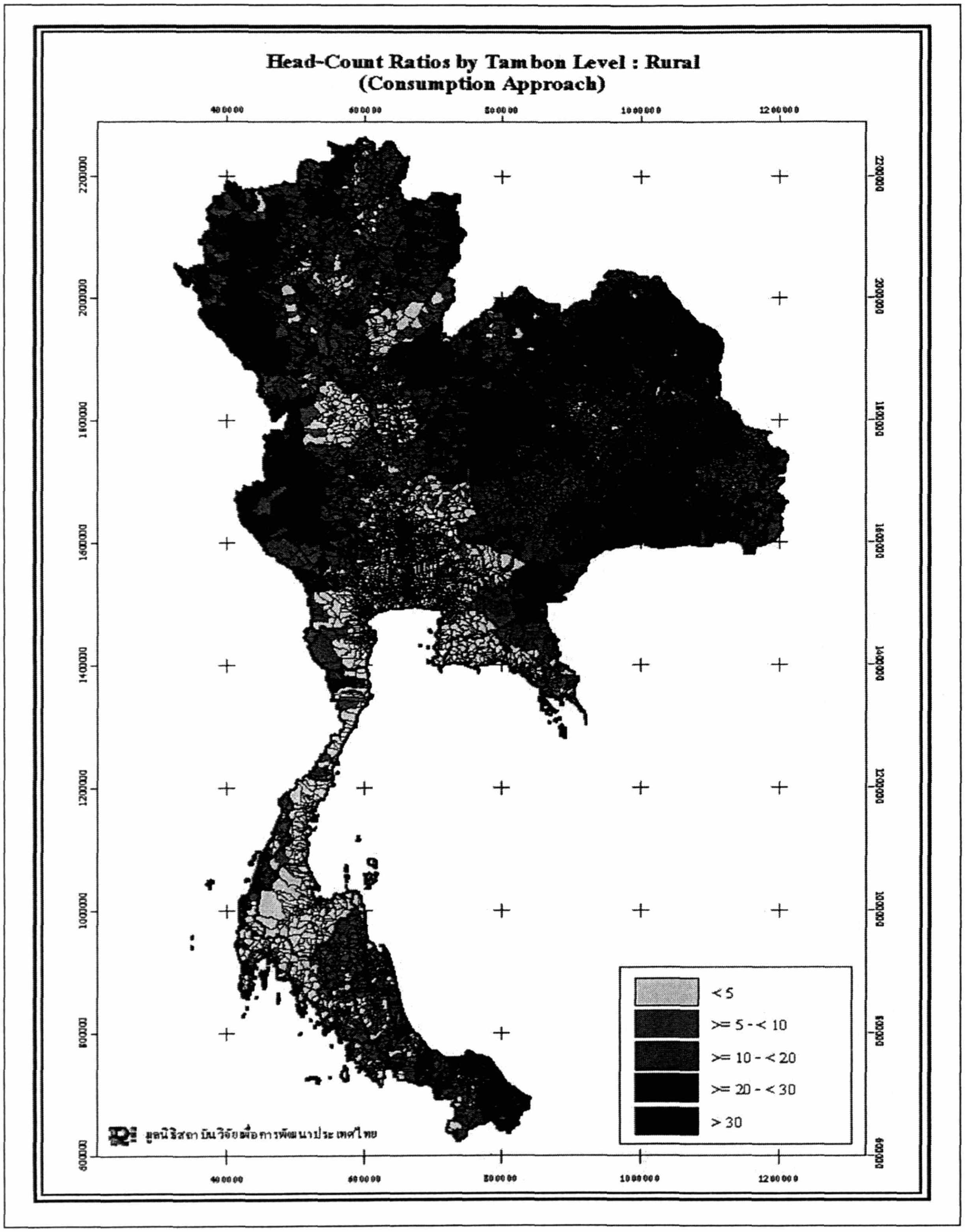

Figure 2: Tambon-level rural headcount for consumption

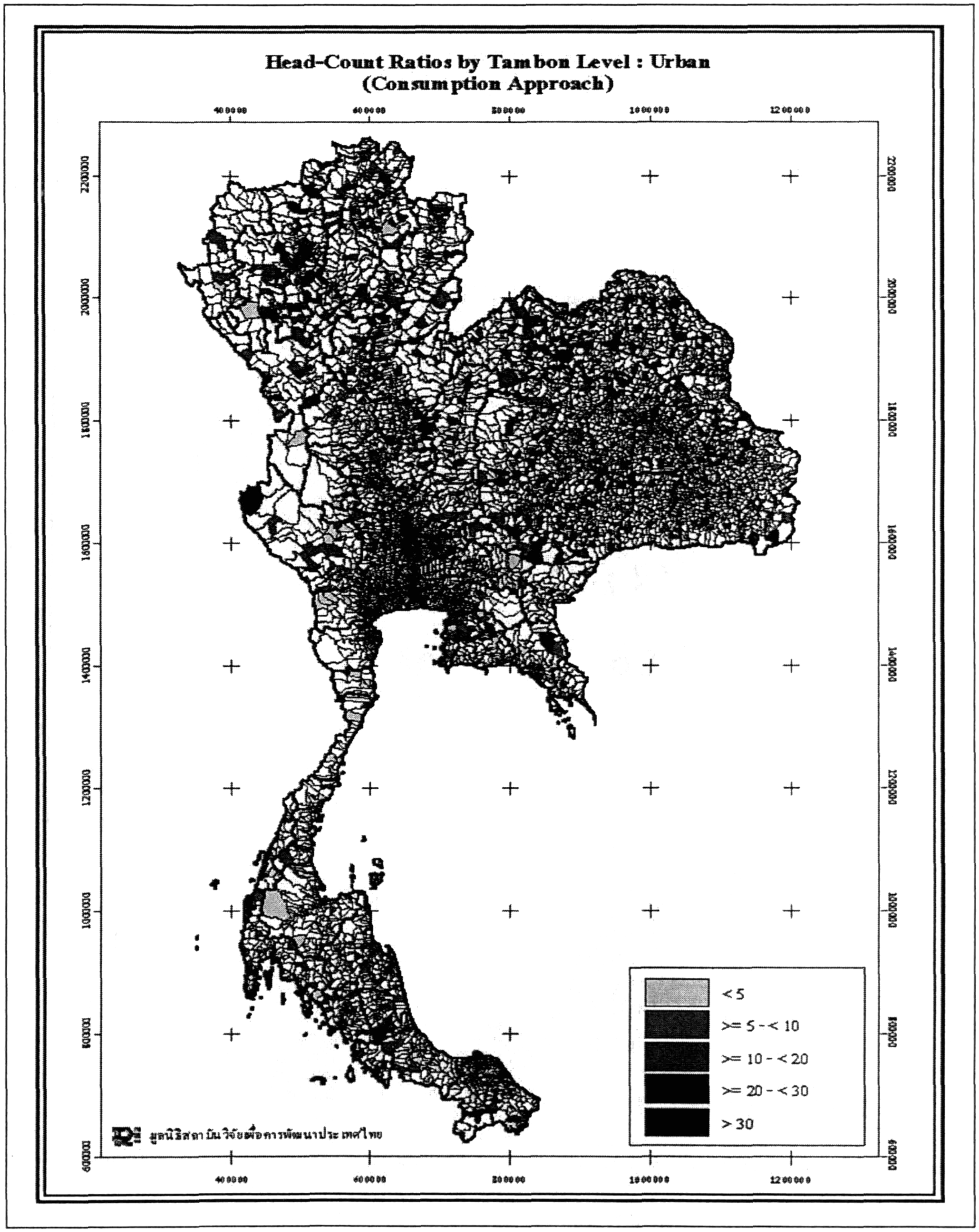

Figure 3: Tambon-level urban headcount for consumption

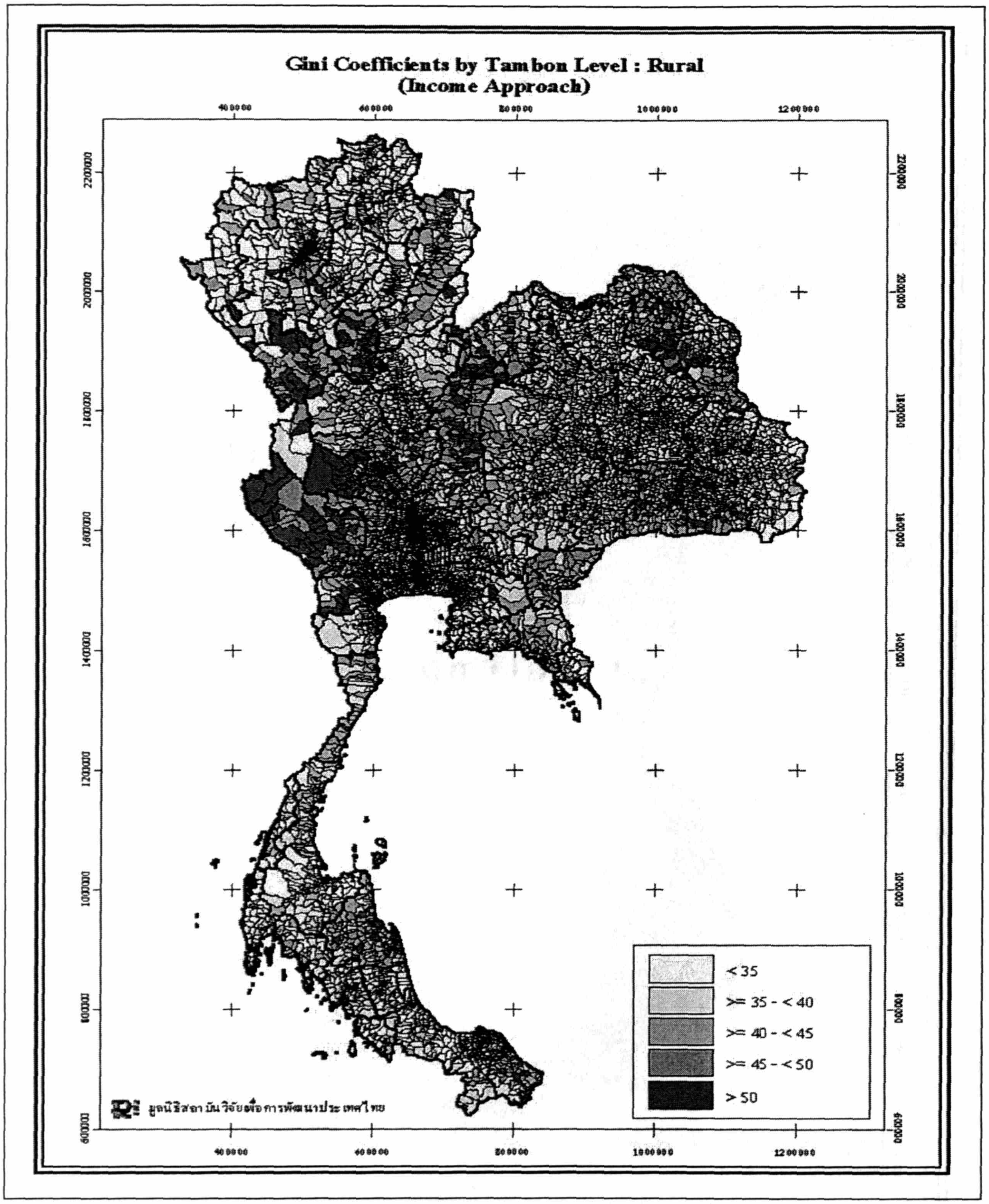

Figure 4: Tambon-level rural Gini for income

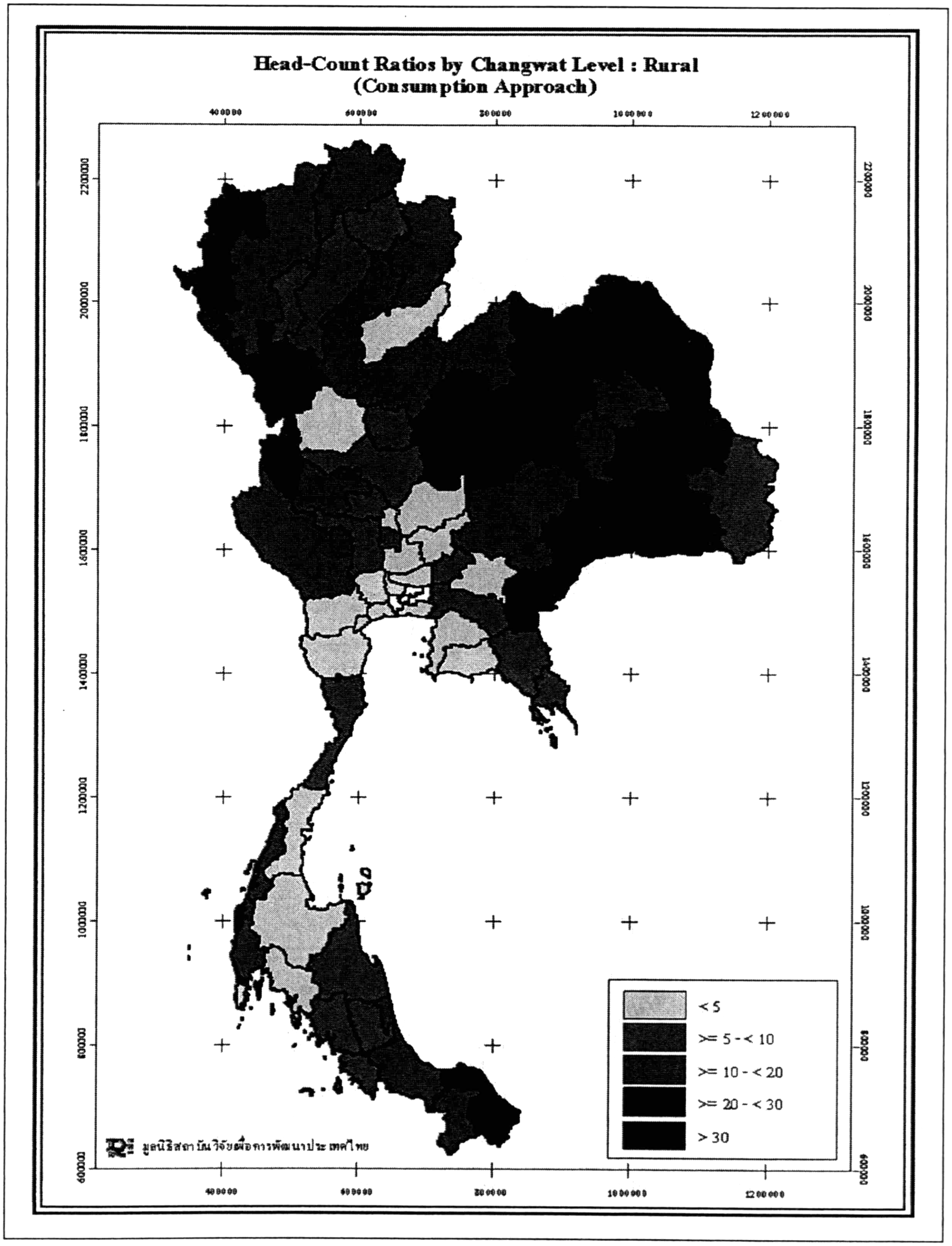

Figure B1: Changwat-level rural headcount for consumption

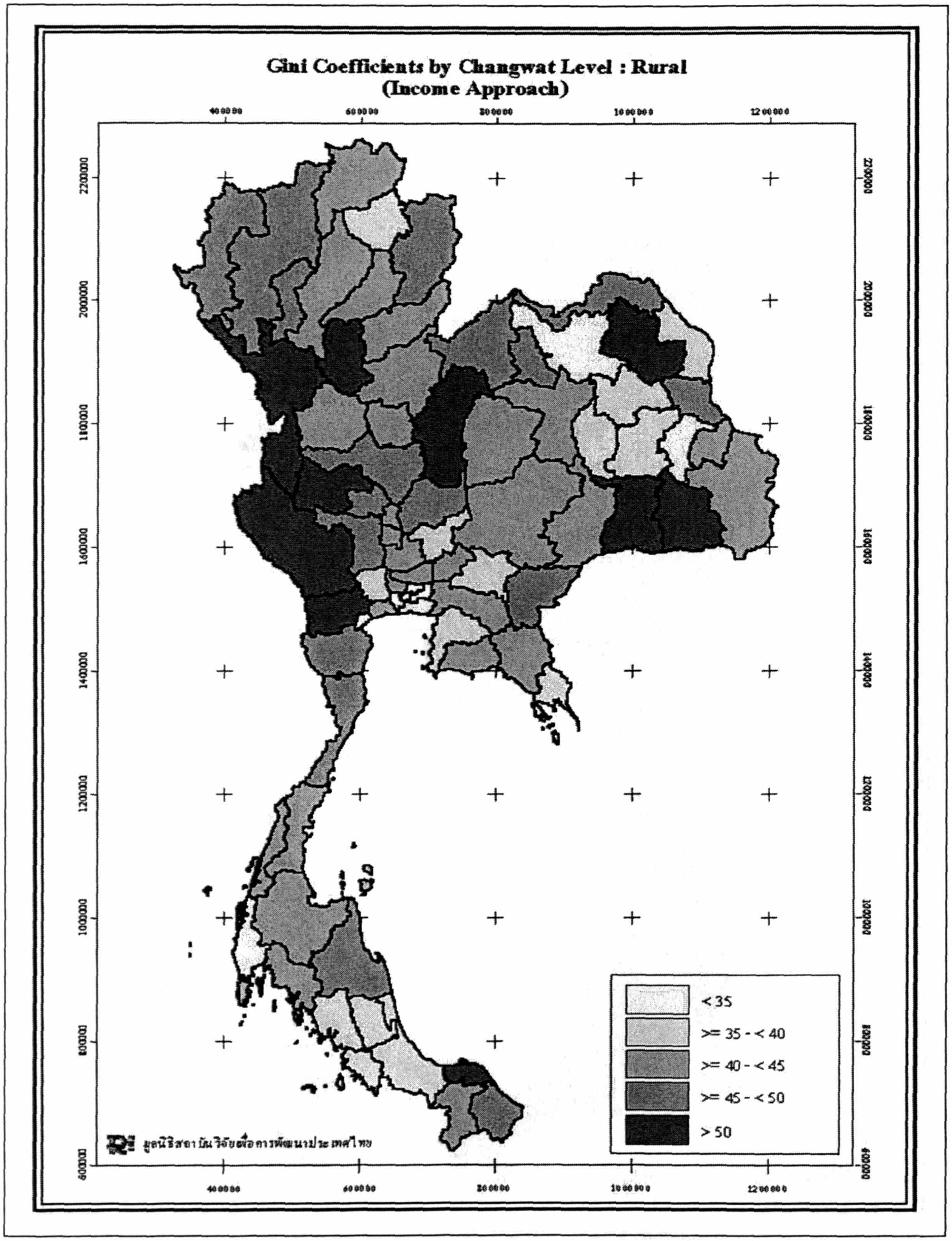

Figure B2: Changwat-level rural Gini for income

Table 1: Hausman tests of parameter equality for rural consumption

Region	Fraction of failures	Median p-value
All (75 changwat)	58.6%	0.024
Central (25 changwat)	60.0%	0.015
North (17 changwat)	47.1%	0.059
Northeast (19 changwat)	73.7%	0.011
South (14 changwat)	50.0%	0.062

Table 2: Quality of model fit for rural consumption

	Income		Consumption	
Region	Median number of regressors	Median number of households	Median R-squared	Average R-squared
All (75 changwat)	11	147	0.712	0.709
Central (25 changwat)	11	141	0.702	0.681
North (17 changwat)	10	153	0.664	0.677
Northeast (19 changwat)	15	161	0.736	0.746
South (14 changwat)	10.5	111.5	0.782	0.759

Table 3: Rural headcount in the map and the SES

	Income		Consumption	
Region	SES headcount	Map headcount	SES headcount	Map headcount
All	19.2	19.3	17	16.2
	(.5)	(.5)	(.7)	(.5)
Central	6.4	8.5	4	6.2
	(1.7)	(1.7)	(1.4)	(1.7)
North	14.7	16.2	12.8	12.9
	(2.1)	(1.5)	(2.5)	(1.5)
Northeast	31.4	29.1	28.5	25.5
	(2.9)	(2.2)	(3.2)	(2.4)
South	12.4	13.5	11.8	11.3
	(3.7)	(2.8)	(3.4)	(3)

Standard errors in parentheses

Table 4: Urban headcount in the map and the SES

	Income		Consumption	
Region	SES headcount	Map headcount	SES headcount	Map headcount
All	8.8	9.4	8.4	9.7
	(.4)	(.4)	(.4)	(.3)
Central	2.9	4	2.7	4
	(.9)	(.8)	(.9)	(.9)
North	8.3	10.4	7.6	11.3
	(1.8)	(1.5)	(1.5)	(1.1)
Northeast	18.4	17.1	18	17.6
	(3)	(2)	(3)	(2)
South	7	7.6	6.2	7.5
	(2.2)	(1.5)	(2.3)	(2)
Bangkok	0.3	0.3	0.3	0.5
	(0.2)	(0.2)	(0.2)	(0.2)

Standard errors in parentheses

Table 5: Rural average income and consumption in the map and the SES

	Income		Consumption	
Region	SES average	Map average	SES average	Map average
All	2347.4	2530.1	1781.8	1925.1
	(32.4)	(33.8)	(22.8)	(28.2)
Central	3568.2	3921.5	2573.8	2833
	(231.1)	(223)	(173.8)	(193.4)
North	2207.5	2358.2	1707.5	1797.9
	(121.9)	(107.6)	(97.1)	(80.3)
Northeast	1678.5	1807.5	1335.8	1433.3
	(87.1)	(67.6)	(50.1)	(41.6)
South	2477.3	2587.2	1874.2	2041.4
	(180.4)	(182)	(116.5)	(116.8)

Standard errors in parentheses

Table 6: Urban average income and consumption in the map and the SES

	Income		Consumption	
Region	SES average	Map average	SES average	Map average
All	4644.9	5320.9	3194.2	3539.5
	(84.2)	(77.1)	(65.6)	(64)
Central	5343.4	6167.5	3817.8	4352.7
	(361.1)	(286.1)	(266.4)	(267.1)
North	4156.4	4805.2	2877	2992.9
	(245.2)	(253.8)	(154)	(115.6)
Northeast	3849.8	4474.9	2481.1	2769.1
	(310)	(293.3)	(158.7)	(160.1)
South	4919.6	5338.1	3297	3541.5
	(316.8)	(257.8)	(226.3)	(174.9)
Bangkok	8206.5	8407.9	5329.9	5466.1
	(426.3)	(595.8)	(184.1)	(257.6)

Standard errors in parentheses

Table 7: Rural Gini coefficient in the map and the SES

	Income		Consumption	
Region	SES Gini	Map Gini	SES Gini	Map Gini
All	42.9	43.7	31.7	32.3
	(.3)	(.3)	(.3)	(.3)
Central	42.8	43.9	31.7	33.4
	(1.8)	(1.4)	(1)	(1)
North	43.9	45.8	33.6	34.6
	(1.2)	(1.3)	(1.4)	(1.2)
Northeast	42.5	42.8	30.3	30.3
	(1.3)	(1.2)	(1.3)	(1.3)
South	42.8	42.2	32.5	32.8
	(1.9)	(1.5)	(1.8)	(1.8)

Standard errors in parentheses

Table 8: Urban Gini coefficient in the map and the SES

	Income		Consumption	
Region	SES Gini	Map Gini	SES Gini	Map Gini
All	45.3	46.3	36.4	37.4
	(.3)	(.4)	(.3)	(.3)
Central	41.9	42.1	34.8	36
	(1.2)	(1.3)	(1.4)	(1.2)
North	46	48.8	37.5	38.6
	(1.1)	(1.3)	(1.2)	(1)
Northeast	50.2	51.3	38.9	39.3
	(2)	(1.6)	(1.7)	(1.5)
South	44.4	45.1	34.5	35.9
	(2)	(1.5)	(1.6)	(1.5)
Bangkok	41.7	40.0	33.0	32.8
	(1.7)	(2.6)	1.1	(2.1)

Standard errors in parentheses

Table 9: Average standard errors in the SES compared to the map in rural areas

Variable	SES Gini	Map Gini	SES Gini
Headcount for income	3.8	1.6	3.3
Headcount for consumption	3.4	1.4	2.9
Mean income	320.7	157.1	265.4
Mean consumption	180.5	79.4	135.9
Income Gini	3.6	2.2	3.2
Consumption Gini	3.2	1.6	2.3

Table 10: Median standard errors in the SES compared to the map in rural areas

Variable	SES Gini	Map Gini	SES Gini
Headcount for income	3.7	1.6	3.1
Headcount for consumption	3.4	1.3	2.7
Mean income	259.5	133.3	232.7
Mean consumption	143.7	60.8	110.8
Income Gini	3.4	2	3.1
Consumption Gini	2.8	1.5	2.2

Table 11: Changwat-level headcount for consumption (selected changwat)

Province	SES estimate	Map estimate (changwat model)
Amnatcharoen	34.0	36.6
(rural)	(6.3)	(2.7)
Chachoengsao	**0**	**5.1**
(urban)	**(0)**	**(1.3)**
Lampang	5.9	7.2
(urban)	(2.6)	(1.4)
Lamphun	6.9	6.0
(rural)	(4.0)	(1.0)
Loei	16.0	14.4
(urban)	(6.4)	(2.9)
Phuket	0	0.9
(rural)	(0)	(0.2)
Samutprakan	0	1.2
(rural)	(0)	(0.4)
Satun	7.8	12.6
(urban)	(3.5)	(2.3)

Standard errors in parentheses

Table 12: Amphoe-level consumption headcount (selected amphoe in changwat

Province	SES estimate	Map estimate (changwat model)
Amnatcharoen	36.0	29.1
(rural)	(3.7)	(3.4)
Chachoengsao	**3.1**	**0.7**
(urban)	**(0.8)**	**(0.4)**
Lampang	4.0	3.9
(urban)	(1.0)	(1.1)
Lamphun	2.2	0.3
(rural)	(0.7)	(0.3)
Loei	9.4	14.8
(urban)	(2.3)	(2.4)
Phuket	0.8	0.1
(rural)	(0.2)	(0.1)
Samutprakan	7.5	1.3
(rural)	(2.0)	(11.2)
Satun	9.4	9.2
(urban)	(2.5)	(1.7)

Standard errors in parentheses

Table A1: Rural headcount in the map and the SES (changwat estimates)

		Income		Consumption	
Changwat	Population	SES headcount	Map headcount	SES headcount	Map headcount
samutprakan	301,306	0	.3	0	1.2
		(0)	(.2)	(0)	(.4)
nonthaburi	218,215	0	1.4	0	0
		(0)	(.2)	(0)	(0)
pathumthani	291,383	.4	4.6	0	3.9
		(.4)	(.7)	(0)	(.4)
ayutthaya	379,988	6.7	4.6	.4	1.6
		(3.5)	(.7)	(.4)	(.5)
angthong	174,755	5.6	11.2	2.1	5.2
		(3.3)	(1.3)	(1.5)	(1.2)
lopburi	558,900	7.5	12.4	1.3	3.7
		(2.7)	(1.5)	(1.2)	(1.1)
singburi	131,189	7.5	7.6	1.3	2.1
		(5.5)	(1.4)	(1)	(.5)
chainat	281,486	9.4	13.7	.9	6.8
		(2.6)	(1.5)	(.5)	(1.3)
saraburi	296,585	4.3	3.1	2.7	3.5
		(1.9)	(.5)	(1.2)	(.8)
chonburi	396,018	0	3.1	0	2.5
		(0)	(.8)	(0)	(.1)
rayong	269,968	1	3	2.1	1
		(1)	(.6)	(2.1)	(.3)
chanthaburi	291,284	9.3	8.2	2.8	6.7
		(3)	(1.9)	(1.5)	(1.4)
trat	141,454	16.9	11.3	8.5	6.7
		(5.3)	(2)	(3.2)	(1.3)
chachoengsao	506,896	1.1	4.9	0	6
		(.9)	(1.2)	(0)	(.9)
prachinburi	273,392	2.9	4.2	3.8	3.7
		(2.3)	(.8)	(2.6)	(.9)
nakhonnayok	191,511	5.3	6.5	5	5.5
		(2.7)	(.8)	(2.3)	(1.1)
sakaeo	549,755	26.2	26.1	17.8	25.4
		(4.6)	(2.3)	(5.8)	(3)
ratchaburi	452,465	1.8	3.6	1.7	2.9
		(1.2)	(.8)	(1)	(.7)
kanchanaburi	484,842	12.2	21.5	16.1	17.7
		(3.1)	(1.3)	(4)	(1.6)
suphanburi	595,813	8.2	11.3	7.9	9.3
		(3.3)	(1.7)	(2.5)	(1.1)
nakhonpathom	498,284	1.7	1.7	0	1.2
		(1.2)	(.2)	(0)	(.4)
samutsakhon	227,073	4.8	7.2	.8	2.5
		(2.4)	(.9)	(.7)	(.5)
samutsongkhram	128,481	0	1.5	0	1
		(0)	(.2)	(0)	(.4)

phetchaburi	225,475	4.7	6.5	2.7	1.9
		(2.6)	(1.2)	(1.6)	(.5)
prachuapkhirikhan	235,318	11.3	9.7	7.8	5.6
		(2.7)	(1.3)	(3.4)	(1.1)
chiangmai	941,464	6.3	12.9	7.5	14.7
		(2)	(1.6)	(2.4)	(1.8)
lamphun	287,801	7	8.4	6.9	6
		(5.1)	(1.4)	(4)	(1)
lampang	500,629	17	17.3	15.4	9.5
		(4.7)	(1.5)	(4.3)	(1)
uttaradit	329,361	11.6	12.4	7.7	3.4
		(3.3)	(1.7)	(2.9)	(.6)
phrae	347,173	8.4	8.3	12.6	12
		(2.9)	(1.2)	(3.8)	(1.8)
nan	325,247	22.8	23.5	13.1	18.7
		(5.8)	(2.4)	(4.6)	(1.7)
phayao	339,250	10.1	15.6	3.4	9.6
		(3.1)	(1.5)	(2.2)	(1.5)
chiangrai	801,312	15.5	13.7	22.5	14.8
		(7.3)	(1.2)	(9.9)	(1.4)
maehongson	147,744	30.5	28.3	34.1	27.1
		(7.7)	(1.9)	(7.2)	(1.4)
nakhonsawan	707,159	22.4	19.5	9.6	9.7
		(5.2)	(1.9)	(5)	(1.5)
uthaithani	223,945	22.3	20.1	12.6	16.8
		(5.8)	(1.5)	(5.4)	(2.5)
kamphaengphet	490,602	.8	2.4	1.1	3.6
		(.6)	(.5)	(.9)	(.8)
tak	293,006	27	22.6	24.6	22.4
		(7.3)	(2.1)	(10.5)	(1.7)
sukhothai	408,968	19.5	27.3	7.9	11.9
		(4.2)	(3)	(5.3)	(1.2)
phitsanulok	549,061	17.5	17.2	19.9	15.7
		(3.8)	(1.7)	(5.8)	(1.4)
phichit	396,885	18.4	17.6	6.6	5.3
		(5)	(1.7)	(3.4)	(1.2)
phetchabun	657,788	14.1	21.1	19.6	22.2
		(4)	(1.9)	(4.7)	(2)
nakhonratchasima	1,625,299	24.7	24.7	19.1	14
		(3.9)	(2.1)	(3.5)	(1.7)
buriram	1,071,262	28.5	27.3	25.8	23.4
		(4)	(2.5)	(5)	(3.3)
surin	951,190	43.4	40.1	46	38.4
		(5.6)	(1.4)	(5.1)	(2)
sisaket	1,009,517	29.1	32	32	30
		(4.1)	(3)	(3.6)	(1.9)
ubonratchathani	1,113,152	20.3	17.9	24.4	17.7
		(3.8)	(1.9)	(5)	(1.7)

yasothon	386,339	55.4	51	41.2	30.8
		(6.8)	(2.6)	(7.1)	(2.6)
chaiyaphum	749,503	18.3	24.4	22	35.2
		(4.7)	(3.2)	(4.5)	(1.7)
amnatcharoen	236,099	24.6	30.5	34	36.6
		(5.4)	(2.1)	(6.3)	(2.7)
nongbualamphu	326,504	53.4	42.6	34.4	34.7
		(10.7)	(3)	(7.6)	(4.4)
khonkaen	1,128,250	20.1	19.9	24.3	24.1
		(4.5)	(1.6)	(4.5)	(1.7)
udonthani	846,432	41.2	33.3	55.4	36.2
		(4)	(2.9)	(3.9)	(1.7)
loei	423,419	40.8	32.4	20.4	17.1
		(6.3)	(2.3)	(5.5)	(1.9)
nongkhai	655,156	39.1	42	12.6	21.3
		(4.4)	(2.5)	(3)	(1.2)
mahasarakham	692,451	15.8	12.9	13.1	12.8
		(4.3)	(1.7)	(5.6)	(2.2)
roiet	934,122	24.2	19	21.8	22
		(3.7)	(2.5)	(4.3)	(2.2)
sakonnakhon	597,862	44.8	36.6	13.9	16.6
		(5.1)	(3.5)	(4.4)	(1.3)
kalasin	732,250	43.1	34.4	50.6	40.7
		(6.3)	(2.4)	(5.7)	(2)
nakhonphanom	454,634	52.9	46.6	38.8	33.7
		(4.1)	(1.8)	(5)	(1.9)
mukdahan	225,299	28.2	30.7	26.7	20.8
		(8.5)	(2.4)	(7)	(2.8)
nakhonsithammarat	982,682	14.7	14.3	11.3	8.4
		(2.5)	(1.3)	(2.6)	(1)
krabi	207,412	8.4	12.1	4.4	8.6
		(2.7)	(1.4)	(2)	(1.5)
phangnga	174,654	1.8	3.6	0	3.2
		(1.9)	(1.8)	(0)	(.9)
phuket	126,179	0	.9	0	.9
		(0)	(.2)	(0)	(.2)
suratthani	476,711	3.3	6	6.2	.7
		(1.9)	(1.3)	(3.5)	(.2)
ranong	93,709	13.1	9.6	12	12.9
		(4.7)	(2)	(4.7)	(1.5)
chumphon	305,526	5.6	11	3.9	3.4
		(2.5)	(2.2)	(2.2)	(.3)
songkhla	657,095	4.6	6.7	11.4	12.1
		(2.7)	(1)	(3.7)	(1.6)
satun	163,997	9.7	7.1	3.2	5.8
		(3.7)	(2)	(2.3)	(.6)
trang	403,387	4.6	5	4.7	6.5
		(2.1)	(1.1)	(2)	(.9)
phatthalung	556,542	3.1	8.5	3.6	9.7

		(2.4)	(2.1)	(1.9)	(1.8)
pattani	359,778	28.8	31.8	35.3	29.1
		(5.2)	(1.8)	(6.6)	(2.4)
yala	212,846	33.7	31.5	10.6	10.6
		(7.1)	(2.2)	(3.3)	(1)
narathiwat	384,928	43	36.4	41.5	38.9
		(11.3)	(2.4)	(6.5)	(3.2)

Standard errors in parentheses

Table A2: Urban headcount in the map and the SES (changwat estimates)

		Income		Consumption	
Changwat	Population	SES headcount	Map headcount	SES headcount	Map headcount
samutprakan	634598	1.1	.7	1.1	1.6
		(1.1)	(.2)	(1.1)	(.3)
nonthaburi	525905	0	2.1	0	1
		(0)	(.3)	(0)	(.1)
pathumthani	309602	3.3	2.7	4.8	1
		(3.2)	(.5)	(3.4)	(.5)
ayutthaya	233386	0	1.4	0	1.2
		(0)	(.7)	(0)	(.2)
angthong	71105	12.5	13.9	6.6	9.6
		(4.4)	(2.3)	(5.4)	(1.8)
lopburi	113191	17.9	15.5	13.9	9
		(7.2)	(2.2)	(4.2)	(2)
singburi	67529	5.4	18.1	.9	3.8
		(2.6)	(3.1)	(.9)	(.9)
chainat	44462	7.6	12.5	1.2	8.9
		(4.5)	(3.9)	(1.2)	(2.7)
saraburi	217995	2.4	6.2	2.3	6.6
		(1.4)	(1.4)	(1.6)	(1.7)
chonburi	544840	1.7	1.7	1.7	5.7
		(1.6)	(.7)	(1.6)	(1.7)
rayong	199983	3.8	4.6	2.2	2.8
		(2.6)	(.9)	(2.2)	(.8)
chanthaburi	152418	6	8.6	5.6	8.6
		(2.3)	(1.9)	(3.2)	(1.9)
trat	47689	1.4	8.3	0	1.7
		(1.5)	(2.3)	(0)	(1.5)
chachoengsao	126751	2.2	2.8	0	5.1
		(1.7)	(3.9)	(0)	(1.3)
prachinburi	65583	0	1.2	1.7	1.4
		(0)	(.2)	(1.7)	(.5)
nakhonnayok	29047	2.3	3.1	5.7	2.3
		(1.6)	(1)	(3.2)	(.9)
sakaeo	65352	4.8	4.9	14.6	7.1
		(2.3)	(1.3)	(5.4)	(1.9)
ratchaburi	231564	2.1	7.5	0	2.3

		(1.9)	(1.1)	(0)	(.7)
kanchanaburi	168643	0	3.9	3.6	6.4
		(0)	(1.6)	(1.9)	(1.8)
suphanburi	145515	7.8	8.8	19.2	17.2
		(5.7)	(1.1)	(6.1)	(3.3)
nakhonpathom	209083	0	2	0	.5
		(0)	(.6)	(0)	(.2)
samutsakhon	192044	7.6	2.9	2	6.9
		(4.1)	(.7)	(1.1)	(.8)
samutsongkhram	53233	.2	4.3	.1	1.3
		(.2)	(1.3)	(.1)	(.6)
phetchaburi	150976	3.9	4.6	2.1	4.3
		(2.4)	(.9)	(1.6)	(1.1)
prachuapkhirikhan	148917	5.6	4.2	4.2	3.5
		(4.5)	(1.1)	(3.6)	(.6)
chiangmai	377971	4.2	14.1	5.2	13.3
		(1.3)	(1.3)	(2)	(2.2)
lamphun	96602	13.2	6.6	7.7	11
		(4.4)	(1.2)	(4.7)	(2.3)
lampang	219980	3.9	7.5	5.9	7.2
		(2.2)	(1.6)	(2.6)	(1.4)
uttaradit	97156	9	6.8	5.4	10.6
		(3.6)	(1.9)	(3.7)	(2.2)
phrae	104367	7.1	7.2	11.7	10.7
		(5.1)	(1.5)	(4.4)	(1.6)
nan	60417	9.6	12	7.9	14.6
		(2.8)	(2)	(2.5)	(2.1)
phayao	109654	27.9	23.7	15.9	16.3
		(8.1)	(2)	(6.7)	(1.9)
chiangrai	191058	12	8.8	10.1	10.9
		(4.4)	(1.2)	(2.9)	(1.1)
maehongson	19842	4.9	5.6	.8	5.7
		(2.9)	(2.2)	(.9)	(2.4)
nakhonsawan	220357	5.7	4.7	5.4	9.7
		(1.6)	(1)	(3.9)	(1.2)
uthaithani	50840	14.1	16.3	3.8	7.8
		(5.9)	(5)	(2.2)	(1.4)
kamphaengphet	84198	0	3.4	2.6	3.6
		(0)	(1.6)	(2.3)	(1.4)
tak	112503	10.9	19.3	12.9	14.4
		(4.8)	(2.8)	(5.6)	(2.5)
sukhothai	112852	20.2	17.5	17.8	14
		(4.6)	(1.5)	(4.2)	(1.4)
phitsanulok	152105	4.4	7	3.6	9
		(1.9)	(2.1)	(1.7)	(1.8)
phichit	113550	.9	4.9	1.2	7.1
		(.6)	(1.1)	(.9)	(1.2)
phetchabun	148842	8.1	10.7	11.5	19.4
		(4.5)	(1.5)	(4.8)	(2)

nakhonratchasi ma	519589	10.1	12.7	10.5	14.7
		(2.7)	(1.5)	(2.5)	(1.5)
buriram	200160	21.2	22.2	22.5	20.6
		(6.1)	(3)	(6.6)	(2.2)
surin	101581	17.6	15.3	19.2	18.6
		(4.6)	(1.4)	(4.2)	(1.4)
sisaket	145457	18.2	16.7	20.3	22.8
		(4.3)	(1.6)	(5.7)	(2.7)
ubonratchathani	263875	8.8	6.3	8.7	7.9
		(3.9)	(1)	(3.4)	(1)
yasothon	60351	17	27.8	21.2	26.9
		(4.6)	(1.3)	(9.2)	(3.2)
chaiyaphum	173745	16.6	16.7	23.2	17.1
		(5.3)	(1.6)	(5.4)	(1.6)
amnatcharoen	62479	17.4	19.9	20	23.1
		(3.4)	(2.2)	(3.7)	(2.6)
nongbualamphu	109137	49.3	40.8	41.7	36.6
		(5.6)	(2.1)	(5.5)	(3.5)
khonkaen	366735	4.9	7.5	6.8	10.9
		(2.3)	(1.1)	(2.6)	(1.3)
udonthani	392029	26.5	16.8	34.3	28
		(6.8)	(1.5)	(7.7)	(1.9)
loei	97889	25.8	18.1	16	14.4
		(2.2)	(1.4)	(6.4)	(2.9)
nongkhai	180612	33.2	27.6	12.4	14.5
		(4.6)	(2)	(4.3)	(2.1)
mahasarakham	114223	2.4	12.2	3.8	4.5
		(1.1)	(1.8)	(2.5)	(1.7)
roiet	150318	21.7	20.4	26.4	23.4
		(6.3)	(2)	(4.6)	(1.5)
sakonnakhon	186188	29.6	26.5	18.4	17.6
		(4.8)	(2.5)	(4.5)	(2.3)
kalasin	142664	23.3	22.4	22.1	19.4
		(7.5)	(1.7)	(8)	(1.1)
nakhonphanom	92652	28.7	21.1	23.4	15.3
		(6.9)	(1.8)	(6.8)	(1.8)
mukdahan	45729	9.8	10.4	5.9	6.2
		(3.2)	(2)	(3.6)	(1.1)
nakhonsithamm arat	254688	11.8	9.4	12.9	10.6
		(4)	(1.1)	(5.2)	(1.2)
krabi	54306	8.9	7.8	7.9	4.5
		(4.9)	(1.3)	(5.4)	(1)
phangnga	32804	6.5	9.4	3.3	7.8
		(2.7)	(1.5)	(2.7)	(1.3)
phuket	87133	0	1.3	0	.6
		(0)	(.5)	(0)	(.4)
suratthani	261910	2.6	5.5	0	1.9

		(1.6)	(.9)	(0)	(.5)
ranong	29879	9.1	6.2	4.7	7
		(3.9)	(1.6)	(2.3)	(1.5)
chumphon	81020	.9	5.9	2.9	2.9
		(.7)	(.9)	(2.3)	(.8)
songkhla	402711	1	2.8	.5	3.9
		(.4)	(.5)	(.5)	(.8)
satun	39523	11.6	11.1	7.8	12.6
		(2.8)	(1.7)	(3.5)	(2.3)
trang	117747	3.6	7.2	2	6.9
		(1.4)	(1)	(1.1)	(1)
phatthalung	70282	7.8	7.8	7.5	6.7
		(2.4)	(1.1)	(2)	(1.2)
pattani	115935	12.2	15.4	16.1	17.5
		(4.1)	(1.8)	(5.5)	(1.7)
yala	113114	21.1	13.5	9.2	5.5
		(6.9)	(1.8)	(4)	(1.2)
narathiwat	161395	15.3	14.2	17.6	21.7
		(4.7)	(1.8)	(4.2)	(2)

Standard errors in parentheses

Table 13: Improvements in reducing the poverty gap achieved by our estimates
(for consumption)

Changwat	Estimated headcount for changwat	Expected improvement in poverty gap reduction by targetting to the amphoe	Expected improvement in poverty gap reduction by targetting to the tambon
Chumphon	0.034	0.017	0.064
Krabi	0.086	0.026	0.188
Nakhonsithammarat	0.084	0.054	0.034
Narathiwat	0.39	0.133	0.06
Pattani	0.292	0.077	0.311
Phangnga	0.032	0.011	0.042
Phatthalung	0.097	0.034	0.056
Phuket	0.009	0.001	0.02
Ranong	0.129	0.196	0.302
Satun	0.058	0.027	0.086
Songkhla	0.121	0.079	0.144
Suratthani	0.007	0.02	0.015
Trang	0.065	0.031	0.056
Yala	0.106	0.073	0.102

Notes: The expected improvement in poverty gap reduction by targetting to the amphoe is the average of the difference between the amphoe headcount and the changwat headcount over the 200 simulations. The expected improvement in poverty gap reduction by targetting to the tambon is the average of the difference between the amphoe headcount and the changwat headcount over the 200 simulations.

Table 14: Improvements in reducing the poverty gap achieved by our estimates
(for income)

Changwat	Estimated headcount for changwat	Expected improvement in poverty gap reduction by targetting to the amphoe	Expected improvement in poverty gap reduction by targetting to the tambon
Chumphon	0.11	0.046	0.098
Krabi	0.121	0.061	0.136
Nakhonsithammarat	0.143	0.124	0.104
Narathiwat	0.364	0.167	0.136
Pattani	0.318	0.074	0.177
Phangnga	0.035	0.066	0.051
Phatthalung	0.085	0.03	0.056
Phuket	0.009	0.003	0.007
Ranong	0.096	0.051	0.228
Satun	0.071	0.026	0.073
Songkhla	0.067	0.088	0.119
Suratthani	0.06	0.037	0.063
Trang	0.051	0.015	0.096
Yala	0.315	0.17	0.055

Notes: The expected improvement in poverty gap reduction by targetting to the amphoe is the average of the difference between the amphoe headcount and the changwat headcount over the 200 simulations. The expected improvement in poverty gap reduction by targetting to the tambon is the average of the difference between the amphoe headcount and the changwat headcount over the 200 simulations.