This paper identifies the determinants of income inequality in Thailand, which increased between 1975 and 1998. Our regression results strongly suggest the significance of agricultural factors. We also obtain some limited evidence that sectoral factors, financial development, and education level disparities play a roughly equally important role in explaining inequality changes in Thailand.

Abstract

This paper identifies the determinants of income inequality in Thailand, which increased between 1975 and 1998. Our regression results strongly suggest the significance of agricultural factors. We also obtain some limited evidence that sectoral factors, financial development, and education level disparities play a roughly equally important role in explaining inequality changes in Thailand.

JEL classification: O15; O53

Keywords: Thailand; Inequality; Agriculture; Financial development

1. Introduction

This paper identifies the determinants of the changes in income inequality that took place in Thailand. Fig. 1 illustrates the recent trend in per-capita GNP for Thailand, which increased from US$ 625 in 1975 to US$ 1831 in 1998 (both in constant 2000 dollars). Consequently, Thailand acquired some status as a more-developed country and was seen as having set a precedent for other developing countries. At the same time, Fig. 2 shows that income inequality in Thailand also increased significantly over the same 24 years. This is significant, because “income inequality is relatively stable within countries,” and it “varies significantly among countries” (Li et al., 1998; p. 26).

* Tel.: +81 6 6368 1121; fax: +81 6 6339 7704.
E-mail address: tmoto@ipcku.kansai-u.ac.jp.

0922-1425/$ – see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.japwor.2004.11.003
While the incidence of poverty in Thailand decreased as a result of its economic growth (Deolalikar, 2002; Kakwani and Krongkaew, 2000), the coincident increase in income inequality remains problematic from the perspective of the policy goal of fairness. In addition, some studies have found significant effects of inequality on economic growth.\(^1\) Therefore, it is important to clarify the determinants of the recent increase in inequality in Thailand.

Kuznets (1955) pointed out the importance of considering the agricultural sector for understanding changes in income distribution during the process of economic development. His insight is even more important because it helps to explain the changes in the distribution of income in Thailand. Studies of income distribution in Thailand have stressed the importance of the role of the agricultural sector (Deaton, 1989; Ikemoto, 1991; Krongkaew et al., 1996). While the share of the agricultural sector in the total GDP decreased (from 27% in 1974 to 12% in 1998), the labor force in the agricultural sector still accounted for 51% of the total labor force in 1998.

We used longitudinal data from 24 years and nine national surveys for 13 regions, taken from Thailand’s Household Socio-Economic Survey (HSES). For developing and more-developed countries, it is not always easy to obtain detailed regional data to identify the determinants of income inequality during development. The Thai HSES is one of the few exceptions. The regression results suggest the significance of agricultural factors, although

---

\(^1\) For example, see Deininger and Squire (1998) and Forbes (2000).
we obtained limited evidence that sectoral factors (which include agricultural factors), financial development and education level disparities played a roughly, equally important role in explaining inequality increases in Thailand. The effects of financial development are of interest, because consolidation of financial markets is on the policy agenda of many developing countries. Note that the effect of financial development on income distribution is theoretically indeterminate. However, the regression results suggest that financial development in Thailand may have tempered the inequality increase.

Deolalikar (2002) analyzed poverty in Thailand, focusing on “poverty incidence”. Deolalikar regressed poverty incidence on income distribution and other determinants, using provincial data, and found that inequality had a negative effect on poverty incidence, both directly and via low growth rates. Using a nonparametric approach, Deaton (1989) showed that higher rice prices benefited Thailand’s rural households in all income brackets, although their study did not consider other determinants of inequality. Ikemoto (1991) provided a detailed, elaborate analysis of income distribution in Thailand during the 1960s, 1970s and 1980s, although the study focused more on a precise description of the changes in income distribution than on clarifying the determinants of these changes via econometric investigation.

Although the insight provided by Kuznets (1955) provoked a substantial number of cross-country empirical studies, there is little agreement as to the determinants of inequality. Many of these studies tested the effect of per-capita GDP on inequality (Papanek and Kyn, 1986; Anand and Kanbur, 1993; Jha, 1996; Osaka, 2000). Others utilized relative variables from the agricultural and nonagricultural sectors (Ahluwalia, 1976; Bourguignon and Morrisson, 1998). Li et al. (1998) did not test Kuznets’ hypothesis directly, but they did try to identify the determinants of inequality by focusing on factors related to political economy and capital market imperfections.

Note, however, that these cross-country analyses had their own set of problems. First, Kuznets’ hypothesis dealt with intertemporal relationships, and so it is not apparent whether this cross-country evidence also applies to a country’s course of development. Second, inequality data are generally not strictly comparable across countries, owing, for
example, to differences in the definitions used (e.g., whether income or expenditure, or household or individual). The “secondary” datasets (e.g., the dataset of Deininger and Squire, 1996) used in most of the cross-country analyses suffer from this same problem.\textsuperscript{2} Third, country-specific determinants of inequality, such as the degree of civil liberty (Li et al., 1998) or trade protection (Bourguignon and Morrisson, 1990), are very difficult to measure and likely to be measured erroneously. Our within-country analysis makes overcoming these problems a priority.

Only a limited number of empirical studies have examined intertemporal inequality changes, largely because of the difficulty in obtaining longitudinal data on the income distributions of developing countries. Estudillo (1997) analyzed data for the Philippines for 1961–1991 and found a significant role for urbanization and education. In addition, Karunaratne (2000) found age to be an important determinant in Sri Lanka between 1963 and 1987. Further, Papanek and Kyn (1986), using a small number of longitudinal observations for 83 countries, found only partial support for the intertemporal Kuznets’ hypothesis. Finally, Deininger and Squire (1998), exploited the longitudinal aspects of their earlier dataset (Deininger and Squire, 1996), and rejected the existence of a contemporaneous link between inequality and income levels.

The remainder of this paper is structured as follows. Section 2 introduces indices of income inequality and presents a broad overview of the changes in income distribution in Thailand by exploiting the decomposability of a distribution index. Section 3 explores the determinants of income inequality and details the empirical framework. Section 4 provides the regression results and Section 5 concludes.

2. Changes in income inequality: an overview

2.1. Decomposability of inequality indices

The average household income in Greater Metropolitan Bangkok was 2.7 times larger than that of the rural areas of the Northeastern Region in 1975–1976. This figure increased to 3.4 in 1998. Today, the prevailing notion is that this increase in inter-regional inequality was the driving force behind the inequality increases that took place in the country as a whole. Therefore, it is important that the actual impact of the increase in inter-regional inequalities on national ones be clarified. To this end, one might wish to decompose the measured inequality of the whole country into its inter- and intra-regional components. Since it is impossible to decompose the Gini Index, which is the most popular of the many inequality indices, we use the mean logarithmic deviation (MLD), which is a decomposable inequality index, in addition to the Gini Index.

The Gini Index is specified by the following equation:

\[
\text{GINI} = \frac{\sum_i \sum_j |y_i - y_j|}{2N^2 \bar{y}}. \tag{1}
\]

This index may be expressed geometrically using the Lorenz Curve.

\textsuperscript{2} Atkinson and Brandolini (2001) discussed various problems with secondary cross-country datasets.
Lambert and Aronson (1993) decomposed the Gini Index into three parts as follows:

\[
\text{GINI} = \text{GINI}_B + \frac{1}{N} \sum_i N_i \text{GINI}_{Wi} + R,
\]

(2)

where \( N \) and \( N_i \) represent the populations of the entire country and of subgroup \( i \), respectively (\( N = \sum_i N_i \)). \( \text{GINI}_B \) denotes the Gini Index of the whole population when the income in each group is equalized perfectly. In other words, this represents the between-group Gini Index. The second term provides the population-weighted average of the Gini Indices of the groups (\( \text{GINI}_{Wi} \)) and may be interpreted as the within-group Gini Index. The interpretation of the last term is somewhat less straightforward. In particular, Lambert and Aronson (1993) claimed that this term is a mixture of between-group and within-group inequality and is not decomposable.

Given this difficulty with the Gini Index, Bourguignon (1979) and Shorrocks (1980) proposed a decomposable inequality index that is defined axiomatically. More specifically, let us consider four axioms that inequality measures ought to satisfy. These include the weak principle of transfers, income scale independence, the principle of population and decomposability. The weak principle of transfers requires that the inequality measure increase as the Lorenz curve goes wholly to the outside. Income scale independence is satisfied when the inequality measure is unaffected by proportional changes in the income of all members of the population. The principle of population stipulates that the inequality measure be independent of changes in the population under constant income shares. Decomposability requires that the inequality of the whole population be a consistent function of the inequality within its subgroups. Any inequality measure that satisfies these four axioms is a generalized entropy measure.³

MLD is one of these generalized inequality measures.

\[
\text{MLD} = \log(\bar{y}) - \frac{1}{N} \sum_i \log(y_i).
\]

(3)

Further, the MLD is decomposable as follows:

\[
\text{MLD} = \text{MLD}_B + \frac{1}{N} \sum_i N_i \text{MLD}_{Wi},
\]

(4)

where the first and the second terms are the between- and within-group MLDs, respectively. The between-group MLD is the MLD of the whole population when the income in each group is equalized perfectly. The within-group MLD provides the population-weighted average of the MLDs of the groups. The advantage of the MLD comes from this intuitively appealing decomposition expressed by (4). Using data from the Household Socio-Economic Survey, we use this index to provide an overview of the general contributors to total inequality by decomposing income inequality in Thailand into its inter- and intra-regional components.

³ More about generalized entropy measures and other inequality indices may be found in Cowell (1995).
2.2. The Household Socio-Economic Survey

The National Statistical Office (NSO) of Thailand has conducted the Household Socio-Economic Survey; since 1968, roughly every 5 years before 1987 and every 2 years thereafter. Its objective has been to collect data on the income, expenditures and other characteristics of households. Data for five regions (the Greater Bangkok Metropolitan Area, and the Central, Northern, Northeastern and Southern Regions) became available after the 1975–1976 survey. Four of the five regions in the HSES have three subregions each. These include the municipal area, the sanitary district and the villages. Only the Greater Bangkok Metropolitan Area has no subdivisions. Therefore, we have data for each of the 13 divisions of the entire country. In particular, we use the published volumes of nine full-scale surveys conducted during from the period 1975–1998 (1975–1976, 1981, 1986, 1988, 1990, 1992, 1994, 1996 and 1998). See the data in Appendix A for more information about the HSES.

We do not use the unit record data released from the NSO only for the surveys conducted from 1986 on, because we wish to emphasize the data for the late 1970s and early 1980s, rather than use larger cross section data. As will be discussed in the next subsection, income disparity increased most rapidly during this earlier period.

Welfare is commonly measured in one of four ways: per-capita income, household income, per-capita expenditure and household expenditure. We use the measure “household income”, which is our only viable option, given that the HSES does not provide precise data on the other three measures consistently.

2.3. Inequality decomposition

In this section, we overview of the changes that took place in the income distribution of Thailand by decomposing the income inequality for the entire country into between- and within-group inequalities using the MLD. We divide the country into 13 subregions and obtain inter- and intra-regional inequalities.

We calculate the country MLD using data from Ikemoto and Uehara (2000). They provided the average household income, by decile group, of households, ordered by household income. These estimates were made from tables in published volumes of the HSES, which provide average household income for households, grouped according to given income brackets. We avoided using this original data to calculate the MLD directly, because the occasional revisions of the income brackets used in the HSES that took place in response to overall increases in nominal income can contaminate the MLD.

Fig. 3 presents the regional decomposition of the country MLD. While the inter-regional inequalities seem much smaller than the intra-regional inequalities, this does not necessarily imply that regional or occupational inequalities are very important. Note that the shares of inter-group inequalities increase as smaller subdivisions are used.

Total inequality increased in the 1980s, and it was relatively stable in the 1990s. Inter-regional inequality increased in the late 1980s and in the early 1990s, and decreased thereafter. This pattern resulted partly from the fact that the Greater Bangkok Metropolitan

---

Footnote 4: The Greater Bangkok Metropolitan Area consists of Bangkok, Nonthaburi, Pahum Thani and Samut Prakan.
Area experienced rapid growth around 1990, and other regions subsequently caught up with Bangkok. Note that the evidence presented in this figure does not support the prevailing notion that the inter-regional inequality increases were the primary impetus for the inequality increases in the entire country.

In the next section, we investigate the determinants of inequality econometrically by increasing the number of observations and by focusing on inequality within each of the 13 regions in Thailand.

3. Empirical framework

Unlike existing cross-country inequality analyses, our country analysis does not need to consider country-specific factors like civil liberty or trade openness. In fact, we only consider five factors: the disparity between the agricultural and nonagricultural sectors, average income, financial services, education level disparities and aging. While keeping in mind the possible omission of some determinants of inequality, we use fixed- and random-effects models in the course of the estimation.

Bourguignon and Morrisson (1998) introduced land distribution variables into their regressions and found them to be significant determinants. However, asset distributions are not included in our regression equation, because asset data are unavailable or incomplete for most periods of the HSES. To partially compensate for this shortcoming, we estimate the model both with and without an AR(1) error term. We use the estimation method of Baltagi and Wu (1999), which provides a feasible GLS procedure for estimating panel regression models with AR(1) disturbances when the dataset is spaced unequally.

The regional distribution data come from the HSES tables entitled “average monthly income per household by per-capita consumption expenditure decile groups”, which means the average total current income of households of each of the decile groups, when

---

5 Financial asset data are unavailable in the HSES. Data on the land distribution of farm operators are available from 1986. The available land distribution data show no sign of large-scale variations or trends.
households are ordered by the per-capita consumption expenditures. MLD and GINI denote the MLD and Gini Index calculated from these tables, respectively. This ordering provides the only option for capturing the movement of regional income inequalities for the period 1975–1998, because more desirable data on “average monthly income per household by per household monthly income decile groups” are not available for the entire period.

We estimate a regression of the form:

\[
(MLD' \text{ or } GINI') = \alpha + \beta_1 (DUAL \text{ or } MLD_B \text{ or } GINI_B) + \beta_2 \log(Y) + \beta_3 \text{FIN}
+ \beta_4 \text{EDU} + \beta_5 \text{AGE} + u.
\]  

(5)

where DUAL, MLD_B and GINI_B are the agricultural and nonagricultural sectors disparity measures that we explain in the next subsection. \( \log(Y) \), FIN, EDU, AGE and \( u \) are the log of average household income, a financial development measure, the education level disparity, average age and an error term, respectively. We expect \( \beta_1, \beta_2, \beta_4 \) and \( \beta_5 \) to be positive. The sign of \( \beta_3 \) is indeterminate theoretically. In the following subsections, we detail each of the inequality determinants.

3.1. The disparity between the agricultural and nonagricultural sectors

Ahluwalia (1976) used the share of agriculture in the GDP as an explanatory variable for inequality. Bourguignon and Morrisson (1998) used the relative labor productivity of the agricultural sector as a more natural regressor. In general, these variables have been used in an attempt to capture the effect of the agricultural sector proposed by Kuznets (1955) directly. The measure of relative labor productivity is defective, however, in that it does not consider the population share of the agricultural sector. For example, the effect of an increase in relative labor productivity is small when the population share of the agricultural sector departs greatly from 1/2. Our modified independent variable, which captures the dualistic nature of the economy, is as follows:

\[
\text{DUAL} = \left(-n_a^2 + n_a\right) \left(\frac{y_{na}}{y_a} - 1\right)^2,
\]

(6)

where \( n_a \) is the regional share of agricultural households, and \( y_{na} \) and \( y_a \) are the regional household income averages of the nonagricultural and agricultural sectors, respectively. Note that DUAL is larger when the population share of the economy is close to 1/2 and the relative household income departs from one.

The HSES data include the average household income for nine occupational groups. Table 1 presents these occupational classifications. Farm operators (who own or rent land primarily) and Farm workers are classified as agricultural households, and the remaining groups are classified as nonagricultural households.\(^6\) Fig. 5 illustrates the changes that took place in the share and relative income of agricultural households. DUAL may be interpreted as an inter-sectoral inequality measure, since DUAL is large under a large household income disparity and equivalent household shares of the

\(^6\) The list of occupational classifications is incomplete for the 1975-1976 survey. See the data in Appendix A for the estimation procedure used to make up for this missing data.
agricultural and nonagricultural sectors. This interpretation leads us to put the MLD or Gini Index between the agricultural and nonagricultural sectors, denoted by MLDB and GINIB, respectively, into the regression equation as alternatives to DUAL.

MLDB and GINIB are the MLD and Gini Index of the regional population, when the income of the two sectors is equalized. Fig. 4 presents the relationship between MLDB and GINIB.

Bourguignon and Morrisson (1998) pointed out that it is tautological to explain the total income distribution using the income distribution between the agricultural and nonagricultural sectors. Therefore, one key aim of our regression analysis is to clarify whether the income distribution between the agricultural and nonagricultural sectors can explain the part of the variation in the total income distribution that cannot be explained by other determinants of income inequality. To this end, we estimate additional regressions of the agriculture–nonagriculture disparity on other determinants of inequality.

---

### Table 1
The occupational classification of HSES

<table>
<thead>
<tr>
<th>Farm operators</th>
<th>Own account, nonfarm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mainly owning land</td>
<td>Entrepreneurs, trade and industry</td>
</tr>
<tr>
<td>Mainly renting land</td>
<td>Professional, technical and administration workers</td>
</tr>
<tr>
<td>Employees</td>
<td></td>
</tr>
<tr>
<td>Professional, technical and administration workers</td>
<td></td>
</tr>
<tr>
<td>Farm workers</td>
<td></td>
</tr>
<tr>
<td>General workers</td>
<td></td>
</tr>
<tr>
<td>Clerical, sales and services workers</td>
<td></td>
</tr>
<tr>
<td>Production workers</td>
<td></td>
</tr>
<tr>
<td>Economically inactive</td>
<td></td>
</tr>
</tbody>
</table>

---

Fig. 4. The Gini Index and the MLD: a numerical example.
3.2. Average income

Most empirical studies of income distribution have included the income level or per-capita GDP as an explanatory variable in their regressions and tested Kuznets’ hypothesis. As Bourguignon and Morrisson (1998) pointed out, the inclusion of such a variable is “only an indirect way of accounting for the dualism” (p. 244). While we include agricultural variables in our regression equation, they may not capture all the distribution changes caused by the sectoral factors of the economy. For example, uneven economic growth between traditional and high-tech industries may have led to increased income inequality. Using a preference assumption that incorporates Engel’s law, Eswaran and Kotwal (1993) argued that technical progress in industrial sectors could only benefit rich people. This also may have lead to increased sectoral inequality. Therefore, our regression equation incorporates household income, along with agricultural variables, as a sectoral determinant of inequality.

As we noted in the Section 1, some studies have found significant effects of distribution on income. The probable interaction between income and inequality creates the possibility of an endogeneity bias in our regression analysis. To avoid this problem, we also estimate the model using instrumental variables for income.

In particular, we use log(Y), or the logarithm of the average monthly household income, as a determinant of income inequality in the regression analysis. In addition, we try to capture the relevance of Kuznets’ inverted U-curve by testing the significance of an additional explanatory variable, the square of log(Y), which has often been found significant in previous studies.  

---

For example, see Jha (1996).
3.3. Financial services

The effect that financial services development may have had on income distribution is not straightforward. Developed financial services enable the poor to borrow from the rich, and this decreases income inequality. However, this type of development also locks in inequality, because all agents can increase their financial assets at the same rate. Conversely, developed financial services are often unavailable to the poor, due, for example, to credit constraints arising from information asymmetries or transaction costs. In this case, financial development can exacerbate income inequality. Therefore, one aim of this study is to examine the direction of the effect that financial services development may have had on income inequality.

Li et al. (1998) used M2/GDP as a measure of financial development in their regression analysis and showed that financial development decreased income inequality. However, their measure is not appropriate for use in our within-country analysis, and other good financial asset data are unavailable. Our proxy for financial development is the ratio of insurance and interest expenditures to income, which is denoted by FIN.

3.4. Education level disparities

We introduce education level disparities as a determinant of income inequality, based on the assumption that more education leads to more income. The HSES provides information about the education level distribution of male household members ages 25 years and older. The standard deviation of education in years, calculated from the distribution, is used as a proxy for education level disparity. The proxy is denoted by EDU. In the course of this calculation, we consider the change in the education system that occurred in Thailand in 1960. As a result of this reform in the education system, the number of years spent in elementary school increased from four to seven years, and those spent in secondary school decreased from six to five years. Table 2 presents details of the years spent to reach each education level. We estimate the ratio of adults taught under the old education system to those taught in the new education system, since some adults participated in each system during most of our sample years. In addition to this complication, most of the HSES tables do not provide the education level distribution for each subregion directly. Therefore, we often need to estimate this from the education level distributions corresponding to each occupation. The data in Appendix A details how we obtained these estimates.

Estudillo (1997) found that increases in the share of household heads with a college degree decreased inequality in the Philippines. Akita et al. (1999) found education level differences were a significant determinant in Indonesia for the period 1987–1993. Bourguignon and Morrisson (1998) used secondary school enrolment as a proxy for the share of skilled workers, but the significance of this variable was not robust over time. Assuming a more educated population could restrain the income growth of the richest

---

8 See Motonishi (2000) for a theoretical analysis of the effect of financial development in a stratified world economy.

9 We ignore the impact of the 1978 education system reform, because the number of adults in the education system after 1978 was very small even in 1998, which is the last year of our sample period.
segment of society, Li et al. (1998) also used secondary schooling as a determinant and obtained a significant result.

3.5. Aging

Deaton and Paxson (1994) pointed out that inequality should grow with age. As people get older, the cumulative effect of chance should diversify their income distribution. As a result, the income inequality of an aging economy should increase. Note that this increase of the inequality need not reflect the diversification of people’s lifetime welfare. Therefore, it is important to distinguish this aging effect from other effects.

Using Japanese household survey data, Ohtake and Saito (1998) showed that half of the increase in the economy-wide consumption inequality during the 1980s could be explained by population aging. Karunaratne (2000) also found a significant effect of aging in Sri Lanka. Conversely, Estudillo (1997) concluded that the contribution of aging was limited in the Philippines.

Fig. 6 shows that the population of Thailand is aging. We use the average age of household heads, AGE, as an explanatory variable, to consider the possible relationship between aging and the income distribution.

![Fig. 6. The changes in age distribution.](source: United Nations Demographic Yearbook)
4. Regression results

We use data from nine surveys conducted between 1975 and 1998. The data comprise a total of 116 observations, which correspond to data for 13 regions times nine surveys, minus one, since one observation was lost due to a defective survey report. For a precise description of the dataset construction process, see the data in Appendix A.

Table 3 compares some of the economic indicators for the 13 regions of our sample and shows their development over the sample period. The most urbanized regions during study period were the Greater Bangkok Metropolitan Area and municipal areas, which were generally marked by an increasing share of households in Thailand, low shares of agricultural households, high average incomes, high average schooling, and young ages of the household heads. The four villages contrasted these urbanized areas in regard to average variable magnitudes, and the four sanitary districts fell in between.

Table 4 reports descriptive statistics for the regressors. Tables 5–7 present the regression results. The regression results obtained using the fixed-effects model (FE) and the random-effects model (RE) are provided in Tables 5a and 5b. Table 5a displays the results from the estimation in which the MLD is used as the dependent variable, and Table 5b presents the same for the Gini Index.

In Table 5a, agricultural factors are captured by DUAL or MLDB. In all the regressions in Table 5a, the p-values of the Hausman test are less than 1%, and the random-effects model is rejected. In regressions 1–4, DUAL is significant. Moreover, log(Y) and EDU are significant and have the expected sign. The effect of FIN is negative, which implies that financial development led to a decrease in income inequality. Conversely, there are no signs of a significant effect of aging.

In regressions 5 and 6, DUAL is replaced by MLDB. These results are very similar to those obtained in regressions 3 and 4. We add AGE as an explanatory variable to regressions 5 and 6 and find that it is insignificant. Therefore, the estimation results are omitted.

Regressions 1–6 in Table 5a are then re-estimated using the Gini Index instead of the MLD as an inequality measure. The results presented in Table 5b are very similar to those in Table 5a.

As a means of capturing the possible effects of Kuznets’ inverted U-curve, we also put the square of log(Y) in all the regressions presented in Tables 5a and 5b and found that it is insignificant. This implies that the data for Thailand provide no evidence of Kuznets’ inverted U-curve.

The insignificance of AGE could be partly due to the fact that the aging process is just starting in Thailand. As we can see from Fig. 6, the most remarkable change in the age distribution was a decrease in the ratio of very young people, some of whom are still not in the labor force, compared with those of more advanced age. In Asia, the effect of aging seems to have been more important for more developed countries like Japan, Korea and Singapore.11

---

10 Since we have at most nine observations for each region, we do not take a unit root approach.
11 More information about the aging process in Asian countries can be obtained from the Asian Development Bank (2002).
<table>
<thead>
<tr>
<th>Variables</th>
<th>Greater bangkok metropolitan area</th>
<th>Central region</th>
<th>Northern region</th>
<th>Northeastern region</th>
<th>Southern region</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Municipal area</td>
<td>Sanitary district</td>
<td>Villages Municipal area</td>
<td>Sanitary district</td>
<td>Villages Municipal area</td>
</tr>
<tr>
<td>Share of household number</td>
<td>1975–1976</td>
<td>12.3%</td>
<td>1.6%</td>
<td>3.3%</td>
<td>14.4%</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>17.5%</td>
<td>2.7%</td>
<td>3.2%</td>
<td>13.5%</td>
</tr>
<tr>
<td>The Gini Index (household income)</td>
<td>1975–1976</td>
<td>0.198</td>
<td>0.208</td>
<td>0.207</td>
<td>0.153</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>0.265</td>
<td>0.187</td>
<td>0.232</td>
<td>0.212</td>
</tr>
<tr>
<td>Share of agricultural households</td>
<td>1975–1976</td>
<td>13.6%</td>
<td>8.4%</td>
<td>47.9%</td>
<td>69.9%</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>2.1%</td>
<td>2.6%</td>
<td>14.7%</td>
<td>30.9%</td>
</tr>
<tr>
<td>Average income of agricultural households (Baht)</td>
<td>1975–1976</td>
<td>2089</td>
<td>1874</td>
<td>2445</td>
<td>2030</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>16548</td>
<td>13406</td>
<td>14244</td>
<td>10861</td>
</tr>
<tr>
<td>Average income of all households (Baht)</td>
<td>1975–1976</td>
<td>3425</td>
<td>3498</td>
<td>2565</td>
<td>2024</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>24673</td>
<td>15154</td>
<td>14792</td>
<td>11240</td>
</tr>
<tr>
<td>Average financial expenditures (Baht)</td>
<td>1975–1976</td>
<td>29</td>
<td>49</td>
<td>26</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>1050</td>
<td>538</td>
<td>562</td>
<td>450</td>
</tr>
<tr>
<td>Average of schooling years</td>
<td>1975</td>
<td>5.43</td>
<td>5.32</td>
<td>4.04</td>
<td>3.68</td>
</tr>
<tr>
<td></td>
<td>1975</td>
<td>10.15</td>
<td>8.12</td>
<td>7.76</td>
<td>7.31</td>
</tr>
<tr>
<td>Standard deviation of schooling years</td>
<td>1975–1976</td>
<td>4.22</td>
<td>4.08</td>
<td>2.89</td>
<td>2.25</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>4.63</td>
<td>4.01</td>
<td>3.82</td>
<td>3.55</td>
</tr>
<tr>
<td>Average age of household head</td>
<td>1975–1976</td>
<td>46.2</td>
<td>46.1</td>
<td>49.2</td>
<td>47.6</td>
</tr>
<tr>
<td></td>
<td>1998</td>
<td>42.5</td>
<td>44.0</td>
<td>47.1</td>
<td>49.8</td>
</tr>
</tbody>
</table>
In order to account for the possible endogeneity bias resulting from the interaction between income levels and their distribution, we also estimated the regressions using instrumental variables (the average age of the household head, the standard deviation of the latter’s age and the average education level) for household income. While the coefficient of income increases slightly from this intervention, the overall results do not change significantly and are omitted.

Table 6 presents the standardized coefficients for the regressions in Tables 5a and 5b. The effect of household income is the largest. Education level disparities have the second largest impact, followed by financial development and agricultural factors.

In order to estimate the effects of each determinants on the inequality increases during the period 1975–1998, we combine the first and second columns in Table 4 and the estimated coefficients of regression 5 in Table 5b. The calculated effects of GINI_B, log(Y), FIN and EDU on the Gini Index are 0.007, 0.061, −0.049 and 0.050, respectively.

While the Durbin–Watson statistics for the fixed-effects estimations in Tables 5a and 5b, do not suggest the presence of a significant positive autocorrelation among the residuals, it is still important to be on the safe side and to consider a possible autocorrelation, since we cannot use asset data. Note that the exclusion of the asset distribution indicator from the set of explanatory variables can lead to the autocorrelation of the residuals, because the asset accumulation of households can cause temporary increases in income inequality to persist in later periods. Since we do not have complete asset data, the assumption of AR(1) disturbances is the second-best policy. We set the lengths of the eight intervals intervening between the nine unequally spaced surveys as follows: 3, 3, 1, 1, 1, 1, 1 and 1.

12 The standardized coefficient is the standard deviation change in the dependent variable caused by a one-standard-deviation change in each independent variable.

13 In addition, these Durbin–Watson statistics are conservative, because there are gaps in our dataset.
Table 5a
Regression results (dependent variable: the MLD)

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Models</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 (FE)</td>
<td>2 (RE)</td>
<td>3 (FE)</td>
<td>4 (RE)</td>
<td>5 (FE)</td>
<td>6 (RE)</td>
</tr>
<tr>
<td>DUAL (the relative household income adjusted by shares)</td>
<td>0.124** (0.036)</td>
<td>0.160** (0.032)</td>
<td>0.117** (0.036)</td>
<td>0.162** (0.032)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MLDₐ (the MLD between the agricultural and nonagricultural sectors)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.779** (0.264)</td>
</tr>
<tr>
<td>log(Y) (the log monthly household income)</td>
<td>0.088** (0.016)</td>
<td>0.036** (0.012)</td>
<td>0.087** (0.016)</td>
<td>0.035** (0.012)</td>
<td>0.090** (0.016)</td>
<td>0.044** (0.012)</td>
</tr>
<tr>
<td>FIN (financial development)</td>
<td>-0.979** (0.304)</td>
<td>-0.243 (0.255)</td>
<td>-1.085** (0.289)</td>
<td>-0.184 (0.226)</td>
<td>-1.085** (0.221)</td>
<td>-0.257 (0.226)</td>
</tr>
<tr>
<td>EDU (the S.D. of schooling years)</td>
<td>0.036** (0.008)</td>
<td>0.020** (0.006)</td>
<td>0.037** (0.008)</td>
<td>0.0210** (0.006)</td>
<td>0.036** (0.008)</td>
<td>0.019** (0.006)</td>
</tr>
<tr>
<td>AGE (the age of head)</td>
<td>-0.002 (0.002)</td>
<td>0.001 (0.001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.335** (0.123)</td>
<td></td>
<td></td>
<td>-0.297** (0.092)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durbin–Watson</td>
<td>1.989</td>
<td>1.125</td>
<td>1.975</td>
<td>1.117</td>
<td>1.901</td>
<td>1.087</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.662</td>
<td>0.382</td>
<td>0.658</td>
<td>0.379</td>
<td>0.651</td>
<td>0.395</td>
</tr>
</tbody>
</table>

Number of observations: 116.
Standard errors are in parentheses.
** Significant at the 5% levels.
Table 5b
Regression results (dependent variable: the Gini Index)

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Models</th>
<th>1 (FE)</th>
<th>2 (RE)</th>
<th>3 (FE)</th>
<th>4 (RE)</th>
<th>5 (FE)</th>
<th>6 (RE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DUAL (the relative household income adjusted by shares)</td>
<td></td>
<td>0.147** (0.047)</td>
<td>0.192** (0.042)</td>
<td>0.138** (0.046)</td>
<td>0.195** (0.042)</td>
<td>0.336** (0.136)</td>
<td>0.569** (0.115)</td>
</tr>
<tr>
<td>GINI_R (the Gini Index between agricultural and nonagricultural sectors)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>log(Y) (the log monthly household income)</td>
<td></td>
<td>0.109** (0.021)</td>
<td>0.041** (0.016)</td>
<td>0.107** (0.021)</td>
<td>0.0410** (0.017)</td>
<td>0.115** (0.021)</td>
<td>0.066** (0.017)</td>
</tr>
<tr>
<td>FIN (financial development)</td>
<td></td>
<td>−1.073** (0.394)</td>
<td>−0.151 (0.333)</td>
<td>−1.209** (0.374)</td>
<td>−0.060 (0.290)</td>
<td>−1.220** (0.290)</td>
<td>−0.233 (0.374)</td>
</tr>
<tr>
<td>EDU (the S.D. of schooling years)</td>
<td></td>
<td>0.050** (0.010)</td>
<td>0.029** (0.008)</td>
<td>0.050** (0.010)</td>
<td>0.028** (0.008)</td>
<td>0.047** (0.010)</td>
<td>0.025** (0.008)</td>
</tr>
<tr>
<td>AGE (the age of head)</td>
<td></td>
<td>−0.002 (0.002)</td>
<td>0.001 (0.002)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td>−0.294** (0.162)</td>
<td></td>
<td>−0.234** (0.121)</td>
<td></td>
<td>−0.465** (0.138)</td>
</tr>
<tr>
<td>Durbin–Watson</td>
<td></td>
<td>2.000</td>
<td>1.039</td>
<td>1.974</td>
<td>1.021</td>
<td>1.928</td>
<td>1.046</td>
</tr>
<tr>
<td>$R^2$-squared</td>
<td></td>
<td>0.659</td>
<td>0.343</td>
<td>0.655</td>
<td>0.6</td>
<td>0.646</td>
<td>0.362</td>
</tr>
</tbody>
</table>

Number of observations: 116.
Standard errors are in parentheses.
* Significant at the 10% levels.
** Significant at the 5% levels.
estimation results using the procedure of Baltagi and Wu (1999) are presented in Table 7. While the effect of MLDB remains significant in this case, the effects of most of the other variables are insignificant. Although we cannot provide conclusive proof of the importance of these estimation results, they do suggest the need for some caution in interpreting Tables 5a, 5b and 6. Specifically, the results in these tables could be contaminated by an autocorrelation of the error terms.

So far, we have treated the agriculture–nonagriculture disparity as a source of economy-wide disparity. However, given the objective of this paper, which is to explore the determinants of income inequality, it is important to explore the determinants of this agriculture–nonagriculture disparity itself.

To this end, we regress the agriculture–nonagriculture disparity on income, the degree of financial development, the standard deviation of years of schooling, and the age of the household head. The results for this estimation are presented in Table 8. The dependent variables are MLDB in regressions 1 and 2, and the income of the nonagricultural sector relative to that of the agricultural sector in regressions 3 and 4. The Hausman test does not reject regression 2, but it does reject regression 4 at the 5% significance level. In contrast to the analysis of economy-wide disparity, the results in Table 8 fail to clarify the determinants of inequality. AGE is a significant determinant of MLDB, but it is not significant for explaining the relative income of the nonagricultural and agricultural sectors. One possible explanation for this inconsistent age effect is that the age of the population has a positive effect on the share of agricultural households.

Other candidates for the determinants of the agriculture–nonagriculture disparity include the prices of agricultural products, harvest conditions, government agricultural policy and the business cycles in the nonagricultural sector. It is very difficult to obtain information of this type for the 24 years and 13 regions of our sample. The overall poor fit of the regressions in Table 8 suggests that the agriculture–nonagriculture disparity provides unique information about economy-wide disparity. Therefore, we believe that it is reasonable to use agriculture–nonagriculture disparity as a determinant of disparity on the economy-wide level.

Table 6
Standardized coefficients

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Regressions</th>
<th>Table 5a</th>
<th>Table 5b</th>
</tr>
</thead>
<tbody>
<tr>
<td>DUAL (the relative household income adjusted by shares)</td>
<td>0.299</td>
<td></td>
<td>0.352</td>
</tr>
<tr>
<td>MLDB (the MLD between the agricultural and nonagricultural sectors)</td>
<td>0.338</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GINIb (the Gini Index between agricultural and nonagricultural sectors)</td>
<td></td>
<td></td>
<td>0.416</td>
</tr>
<tr>
<td>log(Y) (the log monthly household income)</td>
<td>1.055</td>
<td>1.083</td>
<td>1.294</td>
</tr>
<tr>
<td>FIN (financial development)</td>
<td>−0.447</td>
<td>−0.447</td>
<td>−0.497</td>
</tr>
<tr>
<td>EDU (the S.D. of schooling years)</td>
<td>0.653</td>
<td>0.632</td>
<td>0.890</td>
</tr>
</tbody>
</table>

Note: Standardized coefficient denotes standard deviation change in the dependent variable caused by one standard deviation change in independent variable.
Table 7
Regression results (AR(1) disturbances)

<table>
<thead>
<tr>
<th>Independent variables</th>
<th>The MLD</th>
<th>The Gini Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Models</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1 (FE)</td>
<td>2 (RE)</td>
</tr>
<tr>
<td>MLD_b (the MLD between the agricultural and nonagricultural sectors)</td>
<td>1.239** (0.0312)</td>
<td>1.277** (0.210)</td>
</tr>
<tr>
<td>GINI_g (the Gini Index between agricultural and nonagricultural sectors)</td>
<td>0.516** (0.166)</td>
<td>0.663** (0.112)</td>
</tr>
<tr>
<td>Log(Y) (the log monthly household income)</td>
<td>-0.003 (0.007)</td>
<td>0.040** (0.013)</td>
</tr>
<tr>
<td>FIN (financial development)</td>
<td>-0.018 (0315)</td>
<td>0.008 (0.239)</td>
</tr>
<tr>
<td>EDU (the S.D. of schooling years)</td>
<td>0.015 (0.013)</td>
<td>0.010 (0.007)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.015 (0.013)</td>
<td>-0.320** (0.095)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.595</td>
<td>0.431</td>
</tr>
</tbody>
</table>

Number of observations: 116.
Standard errors are in parentheses.
** Significant at the 5% levels.
<table>
<thead>
<tr>
<th>Independent variables</th>
<th>Dependent variables</th>
<th>Models</th>
<th>1 (FE)</th>
<th>2 (RE)</th>
<th>3 (FE)</th>
<th>4 (RE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>log(Y) (the log monthly household income)</td>
<td>Inter-sectoral MLD</td>
<td>0.003 (0.006)</td>
<td>−0.009* (0.005)</td>
<td>0.227 (0.327)</td>
<td>−0.066 (0.229)</td>
<td></td>
</tr>
<tr>
<td>FIN (financial development)</td>
<td></td>
<td>−(0.106) (0.114)</td>
<td>−(0.001) (0.104)</td>
<td>−(8.804) (6.167)</td>
<td>−(0.335) (5.016)</td>
<td></td>
</tr>
<tr>
<td>EDU (the S.D. of schooling years)</td>
<td></td>
<td>0.005** (0.003)</td>
<td>0.004 (0.003)</td>
<td>0.218 (0.152)</td>
<td>0.182 (0.128)</td>
<td></td>
</tr>
<tr>
<td>AGE (the age of head)</td>
<td></td>
<td>0.0014** (0.0006)</td>
<td>0.0015** (0.0005)</td>
<td>0.051 (0.032)</td>
<td>0.023 (0.026)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
<td>0.009 (0.051)</td>
<td>0.009 (0.051)</td>
<td>0.382 (2.422)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durbin–Watson</td>
<td></td>
<td>2.017</td>
<td>0.631</td>
<td>2.108</td>
<td>1.328</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td></td>
<td>0.746</td>
<td>0.174</td>
<td>0.375</td>
<td>0.015</td>
<td></td>
</tr>
</tbody>
</table>

In summary, the effects of the agricultural variable are very robust. The effects of income, financial development, and the education level standard deviation are somewhat sensitive to the model specification of the error term. If we adopt a panel regression without AR(1) disturbances on the grounds of the weak autocorrelation of the residuals, the standardized coefficients suggest that the effects of the variables capturing sectoral factors, financial development, and education level disparities play an important role in explaining Thailand’s changes in income inequality. In terms of 1975–1998 inequality increase, representative regression results show that sectoral factors and education level disparities increased income inequality with the contribution ratio of four to three. Conversely, the coefficient on FIN is negative, which suggests that financial development decreased income inequality. The magnitude of the effect of financial development is almost equal to that of education level disparities.

5. Conclusions

Kuznets (1955) wrote, “What about the trend toward greater inequality due to the shift from the agricultural to the nonagricultural sectors? In view of the importance of industrialization and urbanization in the process of economic growth, their implications for trends in the income distribution should be explored—even though we have neither the necessary data nor a reasonably complete theoretical model”. (p. 12) Although we now have the necessary data and this study reconfirms the importance of sectoral factors, a theoretical model capturing the sectoral factors of a developing economy has yet to be developed in full.

This study shows some limited evidence that the variables related to sectoral shifts in the economy, i.e., the agriculture–nonagriculture disparity and household income, played a significant role in increasing income inequality. While education level disparities also contributed to increasing inequality with somewhat smaller magnitude, financial development had an effect in the opposite direction, which roughly cancels out the effect of education level disparities.

We must admit that there is room for further investigation of the causes of the observed inequality increases in Thailand. Our estimations allowing for AR(1) disturbances suggest that the estimated effects of income, financial development and education level disparities may be exaggerated by the omission of the asset variable from the regression equations. The utilization of precise asset data is left as a future research task.

Acknowledgements

I would like to thank an anonymous referee, the editor (Ryuzo Sato), Anil B. Deolalikar, Yukio Ikemoto, Takehiro Iwaki, Takashi Kihara, Yun-Hwan Kim, Kanokpan Lao-Araya, Tsunao Okumura and seminar participants at the Asian Development Bank, Kyushu University, Nagasaki University, Kyu-Zushi Conference and the Japanese Ministry of Finance. I gratefully acknowledge financial support from the JSPS and the Japanese Ministry of Finance.
Appendix A. Data


The income data are based on “total current income,” which refers to before-tax, after-transfer income, minus “other monetary receipts” (i.e., insurance proceedings, lottery winnings and other receipts).

The three subregions (the municipal area, sanitary district and villages) were designated by the administration according to their degree of urbanization. However, this subdivision was simplified into two categories (the municipal area and villages) before the 2000 survey. For more about the definition of these subregions, see Ikemoto (1991), pp. 31–32.

The HSES volumes for 1975–1976 do not contain information on the percentage of farm workers. By comparing data from 1975–1976 with those from 1981, we can assume that this omission is due to the fact that farm workers were lumped together with general workers. The HSES volumes of the 1975–1976 survey available from the Economic Statistics Division of the NSO include some hand-written corrections. In particular, the writing on Table 5.2 for the 1975–1976 survey in the Northern Region volume also indicates that farm workers were combined with general workers. Therefore, the household ratios and the income of the farm workers in the 1975 survey are calculated, for our purposes, by dividing the data for general workers according to the farm workers/general workers ratio of the 1981 survey.

The HSES volume for the Greater Bangkok Metropolitan Area in 1986 differs from other volumes for the same region. It provides tables for three subregions (the city core, suburbs and fringe areas). However, it does not provide data for the entire region. It also does not provide data of the subregional household numbers, which is necessary for summing subregional data. Therefore, we were unable to obtain income distributions and other data for the Greater Bangkok Metropolitan Area for 1986.

The ratio of adults in the education system before 1960 to those after is assumed to be 1–0 in 1975–1976 survey and 0.31–0.69 in the 1998 survey. The value 0.69 indicates the fraction of adults over 50 years old in 1998 who were 12 years old in 1960. We also assume that the growth rate of the fraction of adults in the new education system was constant from 1975 to 1998.

The HSES volumes for the four regions other than the Greater Bangkok Metropolitan Area in and after 1986 do not provide education level distribution data for the three subregions (the municipal area, sanitary district and villages), while they do provide them for the six occupational groups. Therefore, we estimated the education level distribution for each subregion from the weighted average of the education level distribution for these occupational groups, using the occupational distribution in each subregion as its weight.
The education level classification in the 1975–1976 survey differed from that of the surveys conducted in and after 1981. As a result, we relabeled the “No formal education” class in the 1975–1976 volumes as “Kindergarten or None”. Moreover, no distinction was made between University Education and Higher Education in the 1975–1976 survey. This omission does not present a big problem, because the number of people with “Higher Education”, i.e., with more than a bachelor’s degree, in 1981 was less than 1% in all regions and subregions. Therefore, we assume that the fraction of adults with more than a bachelor’s degree in the 1975–1976 survey equalled zero.

The household head is defined as “the person recognized as such by other members, whether he or she was responsible for financial support or welfare of the household members or not” (HSES 1998).

One may purchase unit record HSES data for 1986 and subsequent years on CD-ROMs from the NSO of Thailand. While some studies have used data from surveys conducted before 1986 “on tape” (Ikemoto, 1991, for example), the availability of the unit record data of these old surveys is uncertain.

More detailed information about the surveys and their design can be found in each of the volumes published by the National Statistical Office (both in Thai and English). See also http://www.nso.go.th/eng/stat/socio/socio.htm.

References


