Methods for producing world and regional estimates for selected key indicators of the labour market

Wes Schaible

Contents

Foreword

The International Labour Office (ILO) plays a major role in assisting member States to promote full employment through economic and social policies that adapt to rapid changes in the world of work. In 1996, the Employment Polices Committee of the International Labour Conference, recognizing the need of governments and the social partners' for timely and accurate information on labour market developments, requested that the ILO develop and disseminate an expanded range of up-to-date and relevant labour market indicators. The Key Indicators of the Labour Market (KILM) Project was designed with two objectives: (1) to develop a set of labour market indicators, and (2) to widen the availability of the indicators to monitor new employment trends.

In 1997, a collaborative effort involving the ILO, experts from the Organization for Economic Cooperation and Development (OECD) and several national statistical offices was undertaken to complete the selection and refinement of the KILM indicators. The indicators were chosen based on three criteria: conceptual relevance, data availability and comparability across countries and regions. The resulting set of 18 indicators was designed to satisfy the ever-increasing demands of governments and the social partners for timely, accurate and accessible information on the world's labour markets."

As the next step in meeting the requirements of our constituents' need for information, the ILO would like to produce world and regional estimates for the following five of its 18 indicators:

Labour force participation rate
Employment-to-population ratio
Employment by sector
Unemployment
Youth unemployment

Production of global estimates would be a straightforward exercise except that at present some countries are able to provide the required data whereas others are not.

Within the ILO two approaches have been considered to address the problem of making world and regional estimates for the desired KILMs. The first approach treats the problem as one of missing data and the second, as a small area estimation problem. This paper presents a number of methods or estimation procedures within each of the two approaches. Characteristics of the methods are mentioned when applicable and general cautions given. The paper also discusses differences and similarities between the two approaches and gives conditions under which particular estimators are identical in each method. The paper concludes with recommendations for a general approach to calculating world and regional estimates given the available data.

Werner Sengenberger
Director
Employment Strategy Department

1. Introduction

1.1 Key Indicators of the Labour Market

As stated in the 1999 Key Indicators of the Labour Market publication, the International Labour Office (ILO) plays a major role in assisting member States to promote full employment through economic and social policies that adapt to rapid changes in the world of work. In 1996, the Employment Polices Committee of the International Labour Conference, recognizing the need of governments and the social partners' for timely and accurate information on labour market developments, requested that the ILO develop and disseminate an expanded range of up-to-date and relevant labour market indicators. The Key Indicators of the Labour Market (KILM) project was designed with two objectives in mind: (1) to develop a set of labour market indicators, and (2) to widen the availability of the indicators to monitor new employment trends.

In 1997, a collaborative effort involving the ILO, experts from the Organization for Economic Cooperation and Development (OECD) and several national statistical offices was undertaken to complete the selection and refinement of the KILM indicators. The indicators were chosen based on three criteria: conceptual relevance, data availability and comparability across countries and regions. The resulting set of 18 indicators was designed to satisfy the ever-increasing demands of governments and the social partners for timely, accurate and accessible information on the world's labour markets.

As a follow-up exercise for the expansion of KILM, the ILO would like to develop procedures for development of world and regional estimates for the following five of these 18 indicators:

Labour force participation rate
Employment-to-population ratio
Employment by sector
Unemployment
Youth unemployment

Development of such estimates would normally be a straightforward exercise, except that, at present, some countries are able to provide the required data and others are not.

1.2 The role of finite population sampling theory

Finite population sampling theory is concerned with making an inference about a defined population from a sample taken from that population. The team wishes to estimate a population parameter using an estimator and the sample data. In addition, the team would also like to provide some measure of the error or at least the uncertainty associated with a given estimator. Perhaps the most common measure of error of an estimator is the mean square error, composed of the sum of the variance of the estimator and the squared bias of the estimator. Biases can rarely be estimated with any degree of confidence. If an estimator is unbiased or approximately unbiased, the variance of the estimator, which can be estimated from the available data, is a satisfactory measure of error of the estimator. The design of the sample, the variability of the variable of interest, the estimator, and the size of the sample all influence the

magnitude of the variance of the estimator. However, as one would expect, given a design, variable, and estimator the variance of the estimator associated with the sample goes to zero when the size of the sample increases to that of the population. Therefore, if one is interested in a population quantity and has appropriate observations on every unit in the population, finite population sampling theory has little to offer. However, this is rarely the case and either resources prohibit the taking of a census or, even if resources are available, all units in the population do not provide data. In addition, there can be definitional problems and issues of measurement error. So even though the objective is to measure every unit in the defined population, there are aspects of this task that are addressed in the survey sampling literature.

Sampling theory is a relatively young field. Smith (1976) reviews its beginnings in the early 1900's and major advancements since. Theoretical work in sampling can generally be grouped into one of two approaches: design-based or model-based. The design-based or randomization-based approach was developed during the 1940's and provided important inferential, design and estimation results. - In the 1970's survey sampling papers using the model-based approach began to appear in the literature. In this approach, the problem of estimating the unknown population total is considered to be one of estimating the unobserved values of the variable of interest, Y. That is, if N is the number of units in the population and we denote units by i, the set of observed units by s, and the set of unobserved units by $\tilde{s}$, then the unknown total, T, to be estimated can be written as: $T = \sum_{i=1}^{N} y_i = \sum_{s} y_i + \sum_{\tilde{s}} y_i$. Since we know $\sum_{s} y_i$, the problem of estimating T is seen to be one of estimating $\sum_{\tilde{s}} y_i$. In the process of estimating the y_i, this approach, among other things, makes clear the role of a model in the argument and is useful in addressing problems where the sampler does not have control of a randomization mechanism. Major contributions have been made using both approaches although most of the recent sampling and estimation literature employ model-based approaches.

1.3   Contents

Within the KILM team, two methods have been considered to address the problem of making world and regional estimates for the desired indicators. Each of these methods is an example of a particular type of estimator. This paper will first present a number of methods within each of the two types. Methods of estimation in the presence of missing data are discussed in Section 2 of this paper. Three important methods are specified and others mentioned briefly. Section 3 contains a similar discussion of indirect or small area estimation methods. Characteristics of the methods are mentioned when applicable and general cautions given. In section 4, the data available for the five indicators of concern here are presented. Attention is given primarily to auxiliary variables and the degree of missing data. The final section discusses differences and similarities between the two types of methods and gives conditions under which particular estimators from each method are identical. The section also contains recommendations for a general approach and a specific method given the available data.

2. Methods of estimation in the presence of missing data

Most results in finite population sampling theory are derived under the assumption that all the required data from a carefully selected sample are available to be used in the estimation process. In sampling practice, however, the surveyor has limited control over the mechanism that determines which sample units provide data, and often only a subset of the required sample data are available for estimation. This lack of control and the resulting possibility of deviation from the design of the data collection effort cause concern that the incomplete sample may be "unrepresentative" or "unbalanced" because characteristics of units that provide data may differ from those that do not. Of even more importance is the possibility that the sample estimate of the variable of interest from the incomplete sample may be in error because this estimate would be different from the estimate from the unobserved, complete sample.

In such practical situations, estimators other than those derived for use with complete samples must be considered. Simmons (1972) clearly states this need. "The first point to recognize is that when data are missing, imputation must take place, either implicitly or explicitly. If the results of the survey are to be used at all, the analyst or consumer is compelled to draw conclusions about the missing evidence. It is not a question of whether to impute, but how." This type of recognition has led to the development and use of a variety of estimators designed to offer protection against incomplete data bias due to unrepresentative or unbalanced incomplete samples. The remainder of this section will discuss some estimators commonly used in such situations.

2.1 Common methods

There are many estimation methods in the literature designed to be used in situations where data collection is subject to non-response or other causes of missing data. Only a number of the more common ones are described here. A more detailed description of the model-based prediction approach applied to the missing data problem and the resulting statistical properties of many of the estimators discussed in this section may be found in Schaible (1983).

2.1.1 *Imputation of the incomplete sample mean*

In the absence of additional information, use of the incomplete sample mean as an estimate of the population mean would seem to be a risky, but possible, way to proceed. The estimator of the population total T is, in general, $\hat{T} = \sum_s Y_i + \sum_{\tilde{s}} \hat{Y}_i$. When the sample mean of the observed units is used to impute for each of the unobserved units, we have

$$\hat{T} = \sum_s Y_i + \sum_{\tilde{s}} \left(\sum_s Y_i \Big/ n_s \right),$$

where n_s is the number of observed units in the sample. When written in this form it is clear that the variable of interest associated with each unobserved unit is estimated by the sample mean of the observed units. It should be noted that this estimator can also be written as $\hat{T} = N\hat{\bar{Y}}$, where, as before, N is the number of units in the population.

For such an estimate to not be misleading, the average value of the variable of interest for the non-responding units must be approximately the same as that for the responding units. More precisely, under the model-based theory, the incomplete sample mean is unbiased as an estimator of the population mean under the model, which specifies that, for each unit in the population, the expected value of the random variable representing the unobserved variable of interest is a constant. That is, $EY_i = \boldsymbol{m}$, for $i = 1,2,\ldots,$N, where N is the number of units in the population.

This assumption that the expected value of the variable of interest for the non-responding units is equal to that of the responding units is specific to each variable of interest and is rarely verifiable. This is not an assumption to be made lightly and is one that a survey practitioner would rarely make unless there is simply no other option.

2.1.2 Poststratification from the incomplete to the complete sample

Post stratification from the complete sample to the population is a technique commonly used in sampling to reduce the variance of an estimator. Cells or post strata are created using one or more variables that are correlated with the variable(s) of interest. For example, many of the national person based surveys conducted in the United States and Canada use age group and sex to create post strata. A "weight" is formed within each post strata that, in the case of simple random sampling for example, is obtained by dividing the number of population units by the number of sample units. The sample size in each post strata is known from the observed sample data. However, the population size in each post strata must be obtained from other sources; a common source is often the country's official statistics. This weight is associated with every sample unit in the post strata. The sum of the weighted sample values provides an estimate of the population total of the variable of interest. An estimate of the mean is obtained by dividing the estimated total by the population size. Even though post stratification is a technique designed to reduce the variance of an estimator, there is no guarantee that it will do so. The reduction, if any, depends on the relationship between the variable of interest and the variable(s) used to create post strata.

In post stratification from the incomplete sample to the complete sample the cells are called weighting or adjustment cells rather than post strata but the principal is the same as above. A "non-response adjustment" weight is formed within each weighting cell, which again is the case of simple random sampling for example, is obtained by dividing the number of complete sample units by the number of incomplete sample units. For each responding unit this non-response adjustment weight is then multiplied by the inverse of the probability of selection to obtain the final sample weight. The sum over the incomplete sample of the weighted values provides an estimate of the population total of the variable of interest. A post stratified estimator of the population total that uses the same post strata to adjust from incomplete sample as to "adjust" from complete sample to population may be written as

$$\hat{T}_p = \sum_{h=1}^{H}\left(\sum_{s_h} Y_{hi} + \sum_{\tilde{s}_h}\left(\sum_{s_h} Y_{hi} / n_h \right) \right),$$

where $h = 1, 2, \ldots,$ H denote post strata and n_h is the number of responding sample units in post strata h. As can be seen from this expression, the post stratified estimator can be viewed as one in which, within each post strata, the sample mean of the observed units is imputed for the unobserved value associated with non-responding units. It should be noted that, in general, the

use of the weighting cell non-response adjustment requires the information to create cells on the complete sample, not on the entire population.

It was mentioned above, that even though post stratification is a technique designed to reduce the variance of an estimator, there is no guarantee that it will do so. Similarly, even though the non-response adjustment using weighting cells is designed to reduce non-response bias, there is absolutely no guarantee that it will achieve that purpose. As above, a reduction, if any, depends on the relationship between the variable of interest and the variable(s) used to create weighting cells.

Under model-based theory, the incomplete sample mean with the weighting cell adjustment is unbiased as an estimator of the population mean under the following model. That is, a model which specifies that, for each population unit within each weighting cell, the expected value of the random variable representing the unobserved variable of interest is a constant, i.e., for post stratum j, $EY_{ji} = \boldsymbol{m}_j$, for $i = 1,2,\ldots,N_j$. Of course, if the variable used to create the post strata is related to Y, this assumption is somewhat more palatable than the one needed for the estimator described in 2.1.1 to be unbiased.

Even though the validity of the model needed for unbiasedness of this method cannot always be tested in practice, it should be noted that this is surely the most common method of non-response adjustment used in population based surveys.

2.1.3 Imputation with a regression model

In the sampling literature devoted to non-response problems in large surveys a distinction is often made between unit and item non-response adjustment methods. When all, or nearly all, of the survey data are missing for one or more sample units, unit non-response methods are considered. Unit non-response methods are characterized by adjusting the sampling weight to account for the unit non-response. On the other hand, when some, but not all, of the survey data is missing, *item* non-response methods are considered for individual variables. Values are imputed into the data record. This distinction arose out of practical limitations and the lack of resources to consider a separate non-response adjustment method for every variable in surveys that might have hundreds of variables. Regression imputation and most of the methods in section 2.1.4 are considered to be item non-response methods. However, for data collections with a small number or variables this distinction is not meaningful and methods of both types should be considered.

Regression imputation is possible when, like the weighting cell adjustment, data are available on the complete sample that are correlated with the variables of interest. However, instead of using this information to create cells and impute the cell mean for missing values, the data are used in a regression model. The incomplete sample data (dependent and independent variables) are used to estimate the parameters of the model and then the missing value (dependent variable) for each unit in the incomplete sample is predicted using the independent variables known for the unit. Although it is not necessary in most applications - for convenience in the application considered in this paper assume the auxiliary variables are available for all non-observed units in the population - the estimator of a population total with regression imputation may be written as

$$\hat{T}_r = \sum_s Y_i + \sum_{\tilde{s}} x_i \hat{\boldsymbol{b}}_s \,,$$

where x_i is the row vector of known auxiliary variables and $\hat{\boldsymbol{b}}_s$ is the column vector of the usual best linear unbiased estimators of regression model parameters.

This estimator that uses regression imputation to represent the unobserved values is model unbiased under the regression model used for imputation. However, there is often little or no evidence outside of the sample data that this model is a valid one. The problem of choosing a regression model in prediction theory in finite population sampling has been investigated. Royall and Herson (1973a,b) studied the robustness of certain models when all sample units provide data and the selection of sample units can be controlled. They concluded that the estimators investigated could be made insensitive to certain departures from the models by selecting "balanced" sampled and by stratifying on a size variable. However, although the sampler has control over the original sample, he or she rarely has control over which units provide data. In fact, it is precisely such lack of control, and the resulting suspicion that the incomplete sample may not be balanced that causes concern.

2.1.4. Other methods

The following methods will be briefly mentioned for the sake of completeness even though, currently, none of these methods are likely to be useful in addressing the problem at hand.

- In the method usually referred to as hot deck imputation, missing values of the variable of interest are replaced by values selected from respondents in the current survey. A number of hot deck procedures have been suggested including the selection of a random value from the overall sample, random imputation within classes, and ordering the units on a variable known for all units and imputing from a nearest responding unit. As an example, the imputation within classes method is as follows: Post strata are defined. Within a post stratum, a sample unit with a missing data item is identified. A second sample unit that has the data item and is in the same post stratum is randomly selected and the existing data item from the second unit is then reproduced and linked with the first unit to represent its missing item. The procedure is followed for each sample unit with missing data.

- A similar method termed cold deck imputation uses data from sources other than the current survey, for example, earlier surveys or historical data.

- A procedure called raking has been suggested for use with missing data. Raking consists of a series of adjustments to make data consistent with external macro data, for example, marginal totals

- Another way of dealing with refusal to provide data is the substitution of a unit having the same or similar auxiliary characteristics as the non-responding unit during the data collection process.

- Imputation of the average of late respondents or respondents that required extra effort to produce a response has also been considered as method for adjusting for non-response.

2.2 Characteristics and cautions

Characterization of a fairly diverse set of estimators is difficult. A few characteristics and the problems associated with them are summarized as follows.

- Estimators that use more complicated methods of imputation may not perform better than the simple mean of the incomplete sample. There is no guarantee that non-response bias will be reduced and the variance of the more complicated estimators is generally larger than that of the mean of the incomplete sample. So a more complicated estimator must reduce bias more than enough to account for the increase in variance to justify the use of a more complicated procedure. In practice, biases are quite difficult to estimate. Any reduction in error depends heavily on the validity of the underlying model justifying the more complicated estimator. In most situations, the practitioner is pragmatically forced to settle for a model determined by the limited auxiliary variables that are available. Models based on such expediency instil little confidence in either the producers or consumers of the estimates.

- Because estimators designed to use with missing data are unbiased under restrictive models that are generally thought to not reflect reality, empirical evaluations play a critical role in the decision whether or not to use a particular estimator. The performance of a particular missing data estimator in a given application depends on the variable(s) of interest and their relationship to the auxiliary variables through the underlying model. Generalization from one variable and application to another is difficult so that each application requires a different empirical evaluation. In the implementation of essentially all survey efforts, some units do not respond and therefore if an estimate is to be produced an estimator designed for use with missing data must be considered. The missing data whose unavailability causes the problem are the same data that would be most useful in the evaluation of missing data estimators and their underlying models. In other words, the need for missing data estimators occurs in precisely situations where data are not available for their adequate empirical evaluation.

- Direct imputation methods other than the hot and cold deck methods will decrease the variability of the sample data. One result of this is that if the imputation method is not taken into account in the variance estimation process, the variance estimates can easily be underestimated.

3. Indirect or small area estimation methods

Much of what follows in this section is taken from *Indirect Estimators in U.S. Federal Programs* (Schaible, ed. 1996). Federal statistical agencies produce estimates of a variety of population quantities for both the nation as a whole and for sub-national domains. Domains are commonly defined by demographic and socio-economic variables. However, geographic location is perhaps the single variable used most frequently to define domains. Regions, states, counties, and metropolitan areas are common geographic domains for which estimates are required. Federal agencies use different data systems and estimation methods to produce domain estimates. Those systems are designed within time, cost and other constraints for the purpose of producing national and domain estimates and use standard, direct estimation methods. Sample sizes within these domains are large enough so that direct estimates meet the reliability requirements of the design. However, there are always other domains of interest with

smaller sample sizes. When a domain sample size is too small to make a reliable domain estimate using a direct estimator, a decision must be made whether to produce estimates using an alternative procedure. The alternative estimators considered are those that increase the effective sample size and decrease the variance by using data from other domains and/or time periods through models that assume similarities across domains and/or time periods. These estimators are generally biased, but if the mean square error of the alternative estimator can be demonstrated to be small compared to the variance of the direct estimator, the selection of the alternative estimator may be justified. In extreme situations, there may be no sample units in the domain of interest and, if an estimate is to be produced, an alternative estimator will be required. Indirect or small area estimators have sometimes been used in the situations.

Terms used to describe indirect estimators can be confusing. Increased interest in non-traditional estimators for domain statistics has occurred recently among survey statisticians and, even though the term "small area estimator" is commonly used, uniform terminology has not yet evolved. This term is frequently used because in most applications of these estimators the domains of interest have been geographic areas. However, the word "small" is misleading. It is the small number of sample observations and the resulting large variance of standard direct estimators that is of concern, rather than the size of the population in the area or the size of the area itself. The word "area" is also misleading since these methods may be applied to any arbitrary domain, not just those defined by geographic boundaries. Other terms used to describe these estimators include synthetic, local area, small domain, sub domain, small subgroups, sub provincial, indirect, and model dependent. Survey practitioners sometimes refer to indirect estimators as "model-based" whereas this term is rarely, if ever, used to describe direct estimators. However, direct estimators can be motivated by and justified under models as readily as indirect estimators.

Indirect estimators have been characterized in the Bayesian and empirical Bayes literature as estimators that "borrow strength" by the use of values of the variable of interest from domains other than the domain of interest. This approach can be used to provide a working definition of direct and indirect estimators for a broad class of population quantities including means and totals.

A *direct estimator* uses values of the variable of interest only from the time period of interest and only from units in the domain of interest.

An *indirect estimator* uses values of the variable of interest from a domain and/or time period other than the domain and time period of interest.

Three types of indirect estimators can be identified. A *domain indirect estimator* uses values of the variable of interest from another domain but not from another time period. A *time indirect estimator* uses values of the variable of interest from another time period but not from another domain. An estimator that is both *domain and time indirect* uses values of the variable of interest from another domain and another time period.

Indirect estimators depend on values of he variable of interest from domains and/or time periods other than that of interest. These values are brought into the estimation process through a model that, except in the most trivial case, depends on one or more auxiliary variables that are known for the domain and time period of interest. To the extent that applicable models can be identified and the required auxiliary variables are available, indirect estimators can be created to produce estimates. The availability of auxiliary variables and an appropriate model relating the auxiliary variables to the variable of interest are crucial to the formation of indirect estimators.

3.1   Common methods

A large number of indirect estimators are described in the literature and an attempt to provide a comprehensive list here would not be constructive. Although certain types of these estimators are similar, in certain cases, whether a particular estimator falls under one heading or another can be debated. However, for discussion purposes groups of similar estimators and simple examples of major groups are presented below.

3.1.1 *The universe meant as an estimator for a domain*

Perhaps the simplest example of an indirect estimator is the use of the mean of the entire sample of the universe as the estimator for the mean of a specific domain population, for example, the use of the mean from a national sample as an estimate for population mean of a particular state or province. A similarly simple example would be the use of the mean for a defined area at a previous time as the estimate for the current time. In the first example, we are using values of the variable of interest from a different domain in the estimation process and in the second, we are using values of the variable of interest from a different time period. That is, in the first case, we have a domain indirect estimator and, in the second case, a time indirect estimator. These estimators are unbiased under the same types of restrictive model as needed for the incomplete sample mean (section 2.1.1) to be unbiased. These estimators are discussed further to illustrate additional points in section 3.2.1.

3.1.2 *The synthetic estimator*

Synthetic estimators may be domain indirect, time indirect or domain and time indirect. For example, a domain indirect synthetic estimator for a population total in domain d and time t may be written as

$$\hat{T}_{(syn),d,t} = \sum_{h=1}^{H} N_{dth} \hat{\bar{Y}}_{.th} \text{,}$$

where $h = 1, 2, \ldots, H$ denotes post strata; N_{dth} denotes the number of population units in domain d, time t, and post stratum h; and $\hat{\bar{Y}}_{.th}$ denotes the sample mean across all domains for time t and post stratum h. Under model-based theory, the synthetic estimator is unbiased if $E(Y_{dthi}) = \boldsymbol{m}_{.th}$ for all i, where i denotes units within poststrata.

One of the first applications of the synthetic estimator is described in *Synthetic State Estimates of Disability* (NCHS, 1968). A later application of this method to unemployment and housing is described by Gonazlez and Hoza (1978). Variations of this method are discussed in Purcell and Kish (1979).

It should be noted that in this original version of the synthetic estimator, the observed sample values are "estimated" rather than representing themselves. That is,

$$\hat{T}_{(syn),d,t} = \sum_{h=1}^{H} \left(\sum_{s_{dth}} \hat{\bar{Y}}_{.th} + \sum_{\tilde{s}_{dth}} \hat{\bar{Y}}_{.th} \right) \text{.}$$

On the other hand, the best linear unbiased predictor which uses the observed values to represent themselves is written as

$$\hat{T}'_{(syn),d,t} = \sum_{h=1}^{H}\left(\sum_{s_{dth}} Y_{dthi} + \sum_{\tilde{s}_{dth}} \hat{\bar{Y}}_{.th} \right).$$

This expression is similar to the one given for the post stratified estimator. The difference between the two is that when the post stratified estimator is used to estimate the total for domain d, the quantity used to impute for the unobserved values in each post strata is the sample mean of the observed values in the post strata and the domain, whereas, the synthetic estimator is seen to use the sample mean of the observed values in the post strata but across all domains.

3.1.3 Regression estimator

As illustrated by the references given below, there are many types of indirect regression estimators in the statistical literature. The most common approach is similar to any standard regression estimator. The defining difference is that for an indirect estimator for a given time and domain, one or more of the model parameters are estimated using at least some data from outside the time period and/or the domain of interest whereas a direct regression estimator requires the model parameters be estimated using only data from the domain of interest. The most common type of indirect regression estimator occurs when data from the total sample is used to estimate all model parameters.

Regression estimators may be direct or, like the synthetic estimator, domain indirect, time indirect or domain and time indirect depending on how the parameters are estimated. For example, a domain indirect regression estimator for a population total may be written as

$$\hat{T}'_{(reg),d,t} = \sum_{s_{dt}} Y_{dti} + \sum_{\tilde{s}_{dt}} x_{dti}\, \hat{\boldsymbol{b}}_{.t}$$

where x_{dti} denotes a row vector of known auxiliary variables and $\hat{\boldsymbol{b}}_{.t}$, is the column vector of the usual best linear unbiased estimators of regression model parameters. The regression coefficients are estimated using y values from one or more domains besides d but within the time period t. Although the synthetic estimator is discussed here as a separate type of estimator, it can be written as a special case of a regression estimator where the auxiliary variables are defined to be variables indicating whether or not each unit is in post stratum h or not. If we let Y_{dti} be the variable of interest, then the domain indirect regression estimator above is unbiased when $E(Y_{dti}) = x_{dti}\,\boldsymbol{b}_{.t}$.

Other regression methods are discussed in McCullagh and Zidek (1987) and Purcell and Kish (1979). Both the United States and Canada have investigated and used indirect regression methods to produce postcensal population estimates for small geographic areas (Long, 1996, Verma and Basavarajappa, 1987). In the United States indirect regression estimators have been used to produce estimates of employment and unemployment (Tiller, Brown, and Tupek, 1996); acreage planted in certain crops (Bellow, Graham, and Iwig, 1996); and personal income, annual income and gross product (Bailey, Hazen, and Zabronsky, 1996).

3.1.4 Other methods

Most of the methods below are discussed in a recent review paper, Small Area Estimation: An Appraisal (Ghosh and Rao, 1994).

- Levy (1979) suggested a combination of the synthetic and regression approaches in a regression adjusted synthetic estimator.
- A composite estimator of a population total may be written as
- $$\hat{T}_{(com),d,t} = w_{dt}\hat{T}_1 + (1 - w_{dt})\hat{T}_2,$$

where w_{dt} is a weight, usually between zero and one, and $\hat{T}_1$ and $\hat{T}_2$ are component estimators. Typically, in small area estimation applications, one component estimator is direct and the other is either domain or time indirect. Note that requiring a component estimator to be direct necessitates that at least one observation be available from the domain of interest. Synthetic and indirect regression estimators can be used even if there are no observations from the domain of interest. There are a variety of approaches to defining the weight w_{dt} for the composite estimator. A characteristic common to most approaches is that the weights have considerable variation across geographical domains. This is a result of the fact that the sizes of the domain samples used for the direct estimator often vary dramatically when samples are designed to make estimates for a higher level of geographic aggregation. Applications are often distinguished by different indirect component estimators and different approaches to estimation the composite estimator weight. The composite estimator is unbiased when both $\hat{T}_1$ and $\hat{T}_2$ are unbiased. A variety of methodological approaches lead to estimators that can be written as composite estimators and many recent applications use estimators of this form.

- Component of variance methods, Bayes methods, empirical Bayes methods, Hierarchical Bayes methods, constrained Bayes methods and empirical best linear unbiased prediction methods are among the many different approaches suggested in the literature.

3.2   Characteristics and cautions

3.2.1   Indirect estimator characteristics

The number of indirect estimators in the literature is large and the number and variety is growing. This makes characterization of these estimators a difficult task. However, some of the general characteristics and practical problems associated with their application, are summarized below. The similarity between certain of these characteristics and those of estimators designed for use with missing data should be noted.

Insight into the differences between direct and indirect estimators may be gained by inspecting their underlying models. Notation will be required. Let

$d = 1, 2, \ldots, D$ denote domains,

$t = 1, 2, \ldots, T$ denote time periods,

$i = 1, 2, \ldots, N_{dt}$ denote units in the population at time t and in domain d, and

Y_{dti} denote the variable of interest associated with unit/observation dti.

Expectation model $E(Y_{dti})$	BLUE for the model parameter	BLUP for $\bar{Y}_{dt}$	Type of estimator
$\boldsymbol{m}_{dt}$	$\hat{\bar{Y}}_{dt}$	$(1/N_{dt})\left(\sum_{s_{dt}} Y_{dti} + \sum_{\tilde{s}_{dt}} \hat{\bar{Y}}_{dt}\right)$	Direct
$\boldsymbol{m}_{.t}$	$\hat{\bar{Y}}_{.t}$	$(1/N_{dt})\left(\sum_{s_{dt}} Y_{dti} + \sum_{\tilde{s}_{dt}} \hat{\bar{Y}}_{.t}\right)$	Domain indirect
$\boldsymbol{m}_{d.}$	$\hat{\bar{Y}}_{d.}$	$(1/N_{dt})\left(\sum_{s_{dt}} Y_{dti} + \sum_{\tilde{s}_{dt}} \hat{\bar{Y}}_{d.}\right)$	Time indirect
$\boldsymbol{m}_{..}$	$\hat{\bar{Y}}_{..}$	$(1/N_{dt})\left(\sum_{s_{dt}} Y_{dti} + \sum_{\tilde{s}_{dt}} \hat{\bar{Y}}_{..}\right)$	Domain and time indirect

For a simple family of models, the table above presents expectation models, the best linear unbiased estimators (BLUE) for the model parameter, the best linear unbiased predictors (BLUP) for the domain and time specific population mean and the names of the resulting estimators. This example illustrates several points that aid in the understanding of indirect estimator characteristics and the relationship between direct and indirect estimators.

1. A domain and time specific model defines a family of models. For example, associated with the single parameter, domain and time specific model, $E(Y_{dti}) = \boldsymbol{m}_{dt}$ are three other models. With appropriate independence and variance assumptions, each model leads to a best linear unbiased estimator (BLUE) for the model parameter and a best linear unbiased predictor (BLUP) for the population mean. The domain and time specific model leads to a direct estimator whereas the three remaining models lead to a domain indirect, a time indirect, and a domain and time indirect estimator.

2. If the Y's are independent with constant variance, then the BLUE's for the parameters of the four models in this family are: 1) the sample mean in the domain and time period of interest for the model parameter, $\boldsymbol{m}_{dt}$, 2) the sample mean for the specified time period across all domains for the model parameter, $\boldsymbol{m}_{.t}$, 3) the sample mean for the specified domain across all time periods for the model parameter, $\boldsymbol{m}_{d.}$ and 4) the sample mean across all domains and time periods for the model parameter, $\boldsymbol{m}_{..}$.

3. The objective in finite population estimation problems is not to estimate a model parameter, but rather to estimate the population mean (or total) for a particular domain and time period. Within the domain and time of interest, the BLUP of the population total is obtained by adding the known sum of the values for sampled units to the estimated sum of the unobserved values for the non-sampled units. In this example, the unobserved value associated with each non-sampled unit is estimated by the BLUE for the corresponding model parameter. The BLUP for the population mean is obtained by dividing the predicted total by the number of units in the population.

4. For the domain and time specific model, the BLUE for the model parameter is algebraically equivalent to the BLUP for the finite population parameter. For the remaining models in the family, the BLUE for the model parameter and the BLUP for

the finite population parameter are not the same. In these cases, the BLUE for the model parameter is an unbiased predictor for the finite population mean, but it is not the BLUP.

5. It is straightforward to verify that the direct estimator is robust against model failure in the sense that it is unbiased, not only under the domain and time specific model, but under each of the models in the family. Indirect estimators are not robust in the same sense; each of the indirect estimators in the family is biased under the domain and time specific model. However, the domain indirect and the time indirect estimators are more robust against model failure than the domain and time indirect estimator in the sense that they are unbiased, not only under the model that leads to each estimator, but also under the $E(Y_{dti}) = \boldsymbol{m}_{..}$ model that leads to the domain and time indirect estimator. Without evidence to the contrary, the domain and time specific model will be the most plausible in the family, and the bias of indirect estimators under this model will continue to be a major source of concern surrounding applications of indirect estimators.

6. This simple example can also be used to help understand the importance of keeping the purpose of the analysis in mind when selecting an indirect estimator. Not all indirect estimators will be equally appropriate for a given analysis. For example, if the purpose of the analysis is to make comparisons across domains for a given time period, it would serve no purpose to use the domain indirect estimator above since this estimator would produce essentially the same estimate for every domain. Even though this is an extreme example, the point is clear. Domain indirect estimators are based on models that assume the expectation of the variable of interest is the same across domains with respect to some model parameter. This inconsistency between the purpose of the analysis and the method used to produce estimates will be avoided if a time indirect estimator is utilized. If, instead of making comparisons across domains, the purpose of the analysis is to make comparisons across time periods within a given domain, it may be appropriate to select from among the domain indirect estimators. However, it should be stressed that, in practice, the performance of both domain and time indirect estimators depends on the available information and how accurately the model that incorporates this information depicts the actual application of interest.

In addition to the characteristics illustrated in the example above, there are several other fairly well known characteristics of indirect estimators.

• A domain and time specific model is implicitly assumed to be true when analyses among domains and over time are conducted. From a best linear unbiased prediction point of view, a domain and time specific model leads to a best linear unbiased direct estimator and also defines a family of indirect models which allow strength to be borrowed from other domains and/or time periods. The direct estimator is unbiased, not only under the domain and time specific model, but also under each of the corresponding indirect models. However, the best linear unbiased indirect estimators associated with the indirect models are not unbiased under the original domain and time specific model. This indirect estimator bias under the more plausible domain and time specific model adds to the uneasiness associated with the use of indirect estimators. It is the primary reason that indirect estimators are generally considered only when resources prohibit the use of direct estimators of adequate reliability.

- If the variance of an indirect estimator will be smaller than that of the corresponding direct estimator since the indirect estimator not only incorporates observations of the variable of interest from the domain and time of concern but also from other domains and/or time periods.

- If the stochastic model underlying an indirect estimator is a satisfactory representation of reality, then the mean square error of the indirect estimator will likely be smaller than that of the corresponding direct estimator. However, many indirect estimators require strong model assumptions that may not be satisfied in most applications. If this is the case, then the mean square error of the indirect estimator may in fact be larger than the variance of the direct estimator. Although estimation of variances and (more importantly) mean square errors of indirect estimators has received attention in the literature, the estimation of a meaningful measure of error for a single small area remains a problem.

- Usually the task at hand is to produce estimates for a number of small areas simultaneously. There is considerable empirical evidence suggesting that the size of an error of an indirect estimator depends on the relationship of the area population value and population values of the other areas from which strength is borrowed. For example, the error in an indirect estimate for a small area with a very large population value is likely to be relatively large and negative, so that the estimate is closer to the population values of small areas that are not so large. All indirect estimators do not display this characteristic to the same extent. However, it can lead to poor performance of indirect estimators when the purpose of the analysis is to identify domains with extreme population values, to rank domains or to identify domains that fall above or below some predetermined level.

- Because indirect estimators are unbiased under restrictive models that are generally thought to not reflect reality, empirical evaluations play a critical role in the decision whether to use an indirect estimator or not. The performance of an indirect estimator in a given application depends on the variable(s) of interest and their relationship to the auxiliary variables through the underlying model. Generalization from one application to another is difficult so that each application requires a different empirical evaluation. In practice, indirect estimators are considered for use in situations where data are not available to support the use of direct estimators. The data that, if available, would support the use of direct estimators are the same data that would be most useful in the evaluation of indirect estimators and models. In other words, the need for indirect estimators is the greatest in precisely those situations where data are not available for their adequate empirical evaluation.

3.2.2 Cautions

There is a fundamental problem associated with the application of indirect estimation methods. A truly plausible model would depend on domain and time specific parameters, but indirect estimators are associated with models that contain one or more parameters that do not vary either over domains, time or both. In addition, in most practical applications, the statistician is pragmatically forced to settle for a stochastic model determined by the ancillary variables that are available. Models based on such expediency instil little confidence in either the producers or consumers of the estimates. Consequently, everyone concerned is usually convinced that the estimation process produces biased estimates.

As evidenced by the large and growing literature on indirect estimation methods, numerous researchers have been working on the challenging problems facing those who must produce estimates with inadequate resources. Many authors suggest new approaches or variations of existing approaches, but few give caution about the dangers associated with the use of indirect estimation methods. As a major exception, Kalton (1987) provides us with precise and compelling words of caution.

". . . a cautious approach should be adopted to the use of small area estimates, and especially to their publication by government statistical agencies. When government statistical agencies do produce model-dependent small area estimates, they need to distinguish them clearly from conventional sample-based estimates. … Before small area estimates can be considered fully credible, carefully conducted evaluation studies are needed to check on the adequacy of the model being used. Sometimes model-dependent small area estimators turn out to be of superior quality to sample-based estimators, and this may make them seem attractive. However, the proper criterion for assessing their quality is whether they are sufficiently accurate for the purposes for which they are to be used. In many cases, even though they are better than sample-based estimators, they are subject to too high a level of error to make them acceptable as the basis for policy decisions".

Indirect estimation can be considered when other, more robust alternatives are unavailable, and then only with appropriate caution and in conjunction with substantial research and evaluation. Even after such efforts, neither producers nor users should forget that indirect estimates may not be adequate for the intended purpose.

4.   Data requirements and recommended methods

4.1   Data requirements

Both missing data and small area estimation methods are based on models relating the unknown information to known information either at a macro or micro level. In regression terminology, we are using known independent auxiliary variables in a model to predict unknown dependent variables. How well we are able to predict in a given application depends on the auxiliary variables available and how they are related to the variables to be estimated, in this case, the key indicators of interest. The KILM Team has evaluated some auxiliary variables. Both geography as represented by sub-regions and the Human Development Index based on gross domestic product per capita, educational attainment, (adult literacy and combined primary, secondary and tertiary enrollment), and life expectancy have been evaluated in a report by María Jeria Caceres (1998). In addition, plots at the world level between labour force participation rates and GDP per capita at purchasing power parity show little or no correlation between these two variables. Plots of these two variables for the latest year are shown in Appendix A, figure 1. Additional plots between KILMs 2, 4, and 8 and GDP for the latest year also show little or no relationship between GDP and the variables under consideration with the exception of KILM 4, employment by sector (Appendix A, figures 2-6).

Further investigation of these variables taking region and then gender into consideration results in the same conclusion. Plots between KILMs 1, 2, 4, 8, and 9 and GDP by region (figures 734) and by gender for the latest year are given in Appendix B (figures 35-48). Each

plot shows the simple linear regression and gives the proportion (R^2) of the variation explained by the regression. It should also be noted that the relationships shown on these plots for KILM 4, especially, for agriculture and services, are strong and generally consistent across regions and gender. Table 1 presents the R^2's for these plots. As can be seen, the regressions for KILM 4 fit rather well compared to those for other KILMs. It should also be noted that countries in two regions, Sub-Sahara Africa and Middle East and North Africa, did not provide data for the latest year. Although only plots and R^2's for the latest year are presented in this paper, plots for the years 1990 and 1995 were also produced and showed similar results.

Table 1. **Proportion of variation (R^2) explained by a simple linear regression between the specified KILM and GDP per capita at Purchasing Power Parity by region and by all regions and gender, latest year**

Regions	KILM 1	KILM 2	KILM 4 Agric.	KILM 4 Services	KILM 4 Industry	KILM 8	KILM 9
Developed	.30	.30	.68	.61	.00	.05	.13
Transition	.07	.00	.68	.36	.55	.00	.42
Asia and the Pacific	.29	.19	.78	.77	.19	.13	.22
Latin America and the Caribbean	.38	.16	.11	.04	.39	.03	.06
Sub-Saharan Africa	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Mid. East and North Africa	N/A	N/A	N/A	N/A	N/A	N/A	N/A
All regions							
Males	.05	.00	.55	.50	.16	.06	.07
Females	.00	.13	.30	.40	.09	.06	.08

N/A = not available, no countries responded.

 Recommendation: Continue the search for additional auxiliary variables that are correlated with one of more of the KILM variables to be estimated. Simple plots and other diagnostics can be used to verify possible relationships between the auxiliary variable and the variable of interest.

4.2 Small area or missing data methodology

The central problem addressed in the paper is the one of estimating world and regional quantities from a census of countries when a number of countries do not provide the required data.

The most common application of small area estimation methodologies is somewhat different than the above application. A properly designed and executed sample from a well-defined population is available. Domain estimates as well as population estimates are required.

The sample is of sufficient size and design to produce adequate direct population estimates. But the sample sizes in some or all domains are not large enough for direct estimators to have acceptable variances. Small area estimation methods are used in such situations to make the domain estimates. The population estimator remains a direct estimator. If an existing small area estimation method were to be used in the process of making KILM estimates for the world and regions within this framework, it would need to be assumed that the responding countries were a properly drawn sample of all countries. This, of course, is not the case.

Another way to frame the problem is to consider there to be two "strata" in regional population: one consisting of responding countries and the other, of non-responding countries. We have a complete census of the responding strata and no observations from the non-responding strata at both the population and domain levels. We need to estimate the missing values in the non-responding strata. This, of course, is the problem addressed in this paper and, with a small modification, is the most common way to describe the standard missing data problem. The minor difference is that in most applications, the data collection process starts with a sample rather than a census.

In certain situations, a missing data estimator and corresponding small area estimator can be quite similar or even identical. As above, we can consider there to be two domains, one containing responding units and the other, non-responding units. In this case, the sum of the two domain indirect regression estimators is equal to the missing data regression estimator. This, of course, assumes we use the same regression model in both estimators. Similarly, the sum of the two domain synthetic estimators is equal to the poststratified estimator when 1) the auxiliary variable(s) used to create poststrata for the synthetic estimator crosses all domains rather than being hierarchical within domains and 2) the poststrata are defined in the same manner for both estimators. In addition, one could, in a missing data application, use a model which "borrows strength" in the same way as a small area estimator. That is, the model parameters are estimated using data from outside the "area" of interest. The distinction between these two procedures is usually, but not always clear.

Recommendation: Consider the problem of making KILM estimates for the world and regions when a number of countries do not provide the required data as a missing data problem. Standard methods of weighting or imputing for missing data should be evaluated and, if feasible, used to make KILM estimates.

4.3   Selection of an estimator

The poststratified estimator is the most commonly used method in multipurpose sample surveys to adjust for non-response and missing data. As mentioned previously, numerous surveys use this method in one form or another to adjust for missing data. One advantage of this method is that it is a method that adjusts for all variables under consideration by assigning "weights" representing the missing units to the responding units. However, a corresponding disadvantage is that, in effect, it requires one set of auxiliary variables for all variables being estimated. Individual models might perform better than such a broad approach if auxiliary information related to each variable of interest can be identified. In many applications of poststratification for non-response, for example in the U. S. Current Population Survey and National Health Interview Survey, the non-response adjustment weight in each adjustment cell is restricted to be 2 or below (Massey et al, 1989). In these surveys, the adjustment cells are quite small and the response rates high (90 to 95 per cent). In surveys with larger adjustment cells the adjustment factor limit is lower, for example, the U. S. National Health and Examination Survey uses 1.35. Since the non-response adjustment weight in an adjustment cell is defined to the reciprocal of the response rate in the cell a restriction of 2 (1.35) corresponds

to a response rate of 50 (74) per cent. In general, the response rate is allowed to fall below the limit in some cells, however, any weight over the limit in an adjustment cell is distributed over a larger cell, usually the entire responding sample within some other poststratification strata. When the response rate is less than the limit, this procedure is equivalent to imputing the adjustment cell mean to the non-respondents that correspond to a response rate equal to the limit and imputing the larger cell mean to the remaining non-respondents. This procedure keeps the weights from becoming too large and therefore increasing the variance associated with the estimates. However, this requirement is somewhat subjective and is a hedge against failure of the assumption that the values of the variable of interest for the missing units are "like" those for the units with data. This is an assumption that should not be made without inspection and, to the extent possible, supporting data.

In survey practice, very few persons argue with the premise that the response rate should be high. For example, the U.S. Office of Management and Budget, which has regulatory authority over U.S. surveys, has long had a guideline that surveys should achieve response rates of at least 75 per cent (OMB, 1979). The U.S. Office of Management and Budget also requires that proposed data collections with an expected response rate of less than 75 per cent must provide special justifications. OMB also takes the position that data collection activities having a response rate of less than 50 per cent should be terminated. In 1985, OMB looked at approximately 600 business surveys (predominantly mail surveys) and found that the median response rate for probability sample surveys was about 90 per cent and the average response rate overall was in the 80-85 per cent range.

A non-response rate is a proxy for the non-response bias associated with an estimation procedure. The bias, which is almost never available except in specially designed research studies, is the real issue. It is possible that an estimator for one variable from a sample with 98 per cent response may have a larger non-response bias than the same estimator for another variable from a sample with 60 per cent response. However, without other more directly related information, the non-response rate is a common and valuable measure of the potential non-response bias associated with a particular estimator.

As can be seen in table 2, annual response rates for the specified key indicators are generally low. Response rates for all the key indicators generally increase over the time period 1981 to 1997. Response rates for KILM 4 shows marked improvement beginning in 1990 although response in 1996 and 1997 dropped considerably. With few exceptions, the annual response rate for KILM 8 is better than that of any of the other key indicators

A preliminary inspection of annual response rates indicates that the response rates vary considerably across regions and sub-regions as well as across the key indicators of interest. In many regions (and sub-regions), the data are not adequate to produce annual estimates. In other regions (developed countries), response seems to be consistently high. An alternative to the publication of annual estimates for five key indicators of interest would be to publish only those years (and variables) that meet some minimum response criterion. However, if this approach was followed, not only would some regional estimates be missing at certain points in time, but also, the world estimate would not be possible (if it were to be calculated as the sum of the regional estimates).

The 1990 – 1997 response rates for KILM 4 and KILM 8 are better than those for the corresponding years of other key indicators; perhaps estimates could be provided for these two key indicators until response rates for other key indicators can be improved.

As an alternative to annual estimates, a policy of the production of estimates in five or ten year intervals could be considered. This approach might be particularly appropriate for those KILM's that are or could be included in the Labour Force Projections project developed by the ILO Bureau of Statistics.

Table 2.   Response rates for KILMS 1, 2, 4, 8 and 9 by year, 1981-1997 (male and female for all KILMS, age 15-64 for KILM 1)

Year	KILM 1	KILM 2	KILM 4	KILM 8	KILM 9
1981	.11	.10	.03	.34	.15
1982	.12	.11	.01	.37	.15
1983	.17	.15	.00	.38	.17
1984	.15	.14	.01	.09	.17
1985	.18	.14	.00	.39	.20
1986	.19	.15	.00	.41	.22
1987	.16	.14	.01	.40	.22
1988	.20	.17	.01	.42	.24
1989	.20	.17	.03	.46	.26
1990	.19	.26	.42	.45	.26
1991	.21	.18	.55	.53	.31
1992	.22	.19	.50	.51	.30
1993	.24	.20	.46	.50	.30
1994	.25	.21	.44	.49	.32
1995	.25	.31	.38	.46	.33
1996	.27	.22	.30	.44	.30
1997	.24	.21	.25	.33	.23

Source: ILO: 1999 Key Indicators of the Labour Market (Geneva, 1999))

Recommendation: It would be possible to produce regional and world estimates with the post stratified approach, using sub-regions as the poststrata within regions. When sub-regions have no data, the regional estimate could be used for the countries in the sub-region. However, it seems that for annual regional estimates, sub-regions too often have inadequate (or no) response to generally recommend this approach without further evaluation. Adequate response and appropriate empirical evaluations will be needed for this as well as any methodology.

KILM 1 – Labour force participation rate: This KILM has very low response, however, response varies by region. For example, in the latest year, the Developed countries region had good response whereas the Sub-Saharan Africa and Middle East and North Africa regions had no responses. If the response rates for this year are indicative of those for other years, annual estimation for all regions should not be attempted at this time. The Developed countries region is the only region for which annual estimates should be considered. Poststratification could be considered to produce the Developed countries estimate, but the present sub-region cells will need to be inspected. It seems that some cells may have no response for some years. Alternatives such as publishing annual estimates for some regions or publishing five or ten year estimates could be considered. If a correlated variable is found, this recommendation should be revisited.

KILM 2 – Employment-to-population ratio: This KILM has the lowest response of the five under consideration. It appears that for some years there are no responses in some regions (e.g., Sub-Saharan Africa and Middle East and North Africa). Other regions have extremely low response rates. Like for KILM 1, the Developed countries region is perhaps the only region for which KILM 2 annual estimates should be considered. Poststratification could be considered to produce the Developed countries estimate, but the present sub-region cells will need to be inspected. As above, alternatives to the production of annual estimates for all regions can be considered. Also, the search for correlated variables should continue.

KILM 4 – Employment by sector: This KILM has the best potential for the production of estimates because of the relatively high response rates and the correlation with a known auxiliary variable, GDP. However, even here it appears that estimates cannot be made for all regions on an annual basis. As for other KILMs, the two regions, Sub-Sahara Africa and Middle East and North Africa, generally have inadequate response to make annual estimates. The KILM Team should evaluate a combination of regression and post stratified methods. A number of options are possible.

a) The high response rates in 1990 allow consideration of a regression method as a partial solution to the estimation problem for this variable. Poststratification could be used to adjust for countries that did not respond in 1990.

b) Use the relationship between GDP per capita at purchasing power parity and employment by sector to impute for the missing countries. Estimate regression (or other model) parameters using the data from those countries that do respond for the year of interest. The estimated model parameters can then be used to make predictions for the non-responding countries.

c) Use GDP as a poststratification variable.

The fact that KILM 4 has relatively high response rates in 1980 and 1990 presents an interesting evaluation opportunity to evaluate non-response adjustment methods for this variable. An evaluation of the performance of these approaches should be undertaken. Sure an evaluation will not be definitive, but should help support the selection of a methodology.

KILM 8 - Unemployment: The relatively constant response over time suggests that a regression or post-stratified approach might be considered for this KILM if a correlated variable can be identified.

KILM 9 – Youth unemployment: Response for this KILM is not as good as for KILM 4 or 8. A strategy similar to that for KILM's 1 and 2 can be followed.

This paper often refers to imputed values for missing data. This is a convenient way to conceptualize and discuss missing data methods. Most methods do not explicitly calculate and impute individual values. It is suggested that individual values not be calculated and imputed but that the regional estimates be calculated directly.

It also is suggested that a technical appendix or note should accompany all regional and world estimates. The estimation methodology and KILM response rates should be included in this note.

Efforts to increase the response rates of key data items to more acceptable levels should be continued and, if possible, expanded. The KILM Team does not want simply to produce estimates; but rather to produce estimates that are credible and can withstand scrutiny.

References

Bailey, W., Hazen, L., and Zabronsky, D. 1996. "State, Metropolitan Ares, and County Income Estimation" in Schaible, W.L. (ed.), *Indirect Estimators in U. S. Federal Programs,* Lecture Notes in Statistics, No 108, (New York, Springer).

Bellow, M., Graham, M., and Iwig, W. C. 1996. "County Estimation of Crop Acreage Using Satellite Data", in Schaible, W.L. (ed.), *Indirect Estimators in U. S. Federal Programs,* Lecture Notes in Statistics, No 108, (New York, Springer).

Caceres, M.J. 1998. "Regional and World Aggregate Estimates", Unpublished ILO report.

Ghosh, M. and Rao, J. N. K. 1994. "Small Area Estimation: An Appraisal", *Statistical Science*, Vol.9, No. 1, pp. 55-93.

Gonzalez, M. E. and Hoza, C. 1978. "Small area estimation with application to unemployment and housing estimates", *Journal of the American Statistical Association* 73, pp. 7-15.

Johnson, Lawrence Jeff 1999. "Key Indicators of the Labour Market, Overview", Presented at the Bureau of Labor Statistics, Washington D. C.

Kalton, G. 1987. "Panel Discussion" in Platek R., Rao J.N.K., Sarndal C.E. and Singh M.P. (eds,), *Small Area Statistics*, (New York, John Wiley and Sons).

Key Indicators of the Labour Market, 1999. International Labour Office, Geneva.

Levy, P. S. 1979. "Small Area Estimation – Synthetic and Other Procedures, 1968-1978", in Steinberg, J. (ed.), Synthetic Estimates for Small Areas, NIDA Research Monograph 24, U.S. Government Printing Office, Washington, D.C.

Long, J. F. 1996. "Postcensal Population Estimates: States, Counties, and Places" in Schaible, W.L. (ed.), *Indirect Estimators in U. S. Federal Programs,* Lecture Notes in Statistics, No 108, (New York, Springer).

Massey, J. T., More, T. F., Parsons, V. L., Tadros, W. *National Health Interview Survey, 1984-94, Design and Estimation*. PHS 89-1384, U.S. Government Printing Office, Washington, D.C.

McCullagh, P. and Zidek, J. V. 1987. "Regression Methods and Performance Criteria for Small Area Population Estimation" in Platek R., Rao J.N.K., Sarndal C.E. and Singh M.P. (eds,), *Small Area Statistics*, (New York, John Wiley and Sons).

National Center for Health Statistics 1968. *Synthetic State Estimates of Disabi*lity P.H.S. Publication No. 1759, U.S. Government Printing Office, Washington, D.C.

Office of Management and Budget (US) 1979. Memorandum for Heads of Executive Departments and Agencies, Subject: Fiscal Year 1979 Paperwork Reduction Program, January 10, 1979.

Purcell, N. J. and Kish, L. 1979. "Estimation for Small Domains", *Biometrics*, 35, pp. 365-384.

Royall, R. M. and Herson, J. 1973a. "Robust estimation in finite populations I", *Journal of the American Statistical Association* 68: pp.880-889.

Royall, R. M. and Herson, J. 1973b. "Robust estimation in finite populations II", *Journal of the American Statistical Association* 68: pp.890-893.

Schaible, W. L. 1983. "Estimation of Finite Population Totals from Incomplete Sample Data: Prediction Approach" in *Incomplete Data in Sample Surveys*, Vol. 3, pp. 131-142, Academic Press, Inc.

Schaible, W. L. 1996. "Introduction, Recommendations and Cautions" in Schaible, W.L. (ed.), *Indirect Estimators in U. S. Federal Programs,* Lecture Notes in Statistics, No 108, (New York, Springer).

Simmons, W. R. 1972. *Operational control of sample surveys*. Laboratories for Population Statistics, Series No. 2, The University of North Carolina, Chapel Hill, North Carolina.

Smith, T. M. F. 1976. "The Foundations of Survey Sampling: a Review", *Journal of the Royal Statistical Society*, Series A, No. 139, Part 2, pp. 138-204.

Tiller, R., Brown, S., and Tupek, A. 1996. "Bureau of Labor Statistics' State and Local Area Estimates of Employment and Unemployment" in Schaible, W.L. (ed.), *Indirect Estimators in U. S. Federal Programs,* Lecture Notes in Statistics, No 108, (New York, Springer).

Verma, R. B. P. and Basavarajappa, K. G. 1987. "Recent Developments in the Regression Method for Estimation on Population for Small Areas in Canada" in Platek R., Rao J.N.K., Sarndal C.E. and Singh M.P. (eds,), *Small Area Statistics*, (New York, John Wiley and Sons).

Appendix A

Figures

1) KILM 1: Labour force participation rate of both sexes aged 15 years and over, and GDP per capita at purchasing power parity (PPP), latest year
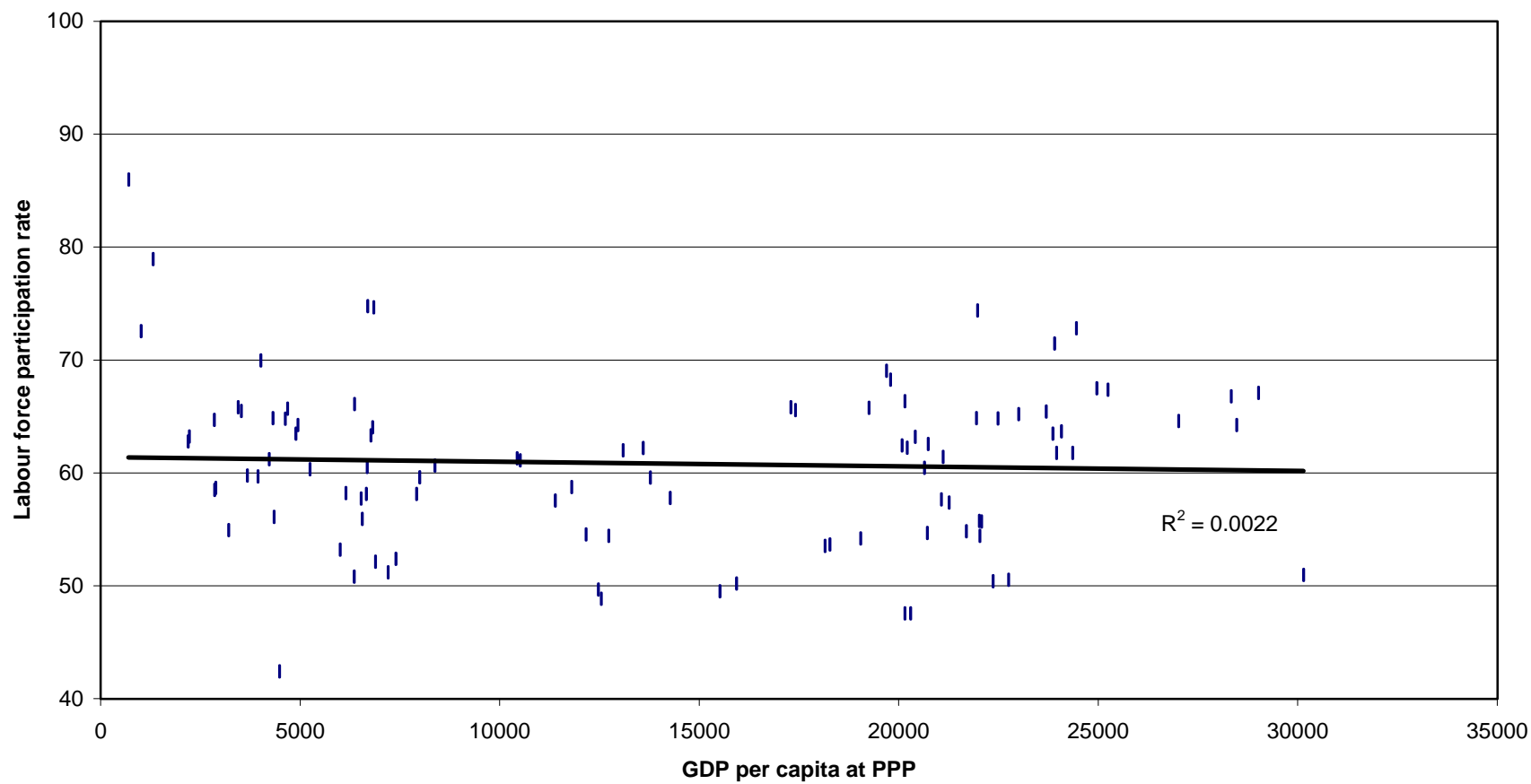2) KILM 2: Employment-to-population ratio of both sexes and GDP per capita at purchasing power parity (PPP), latest year
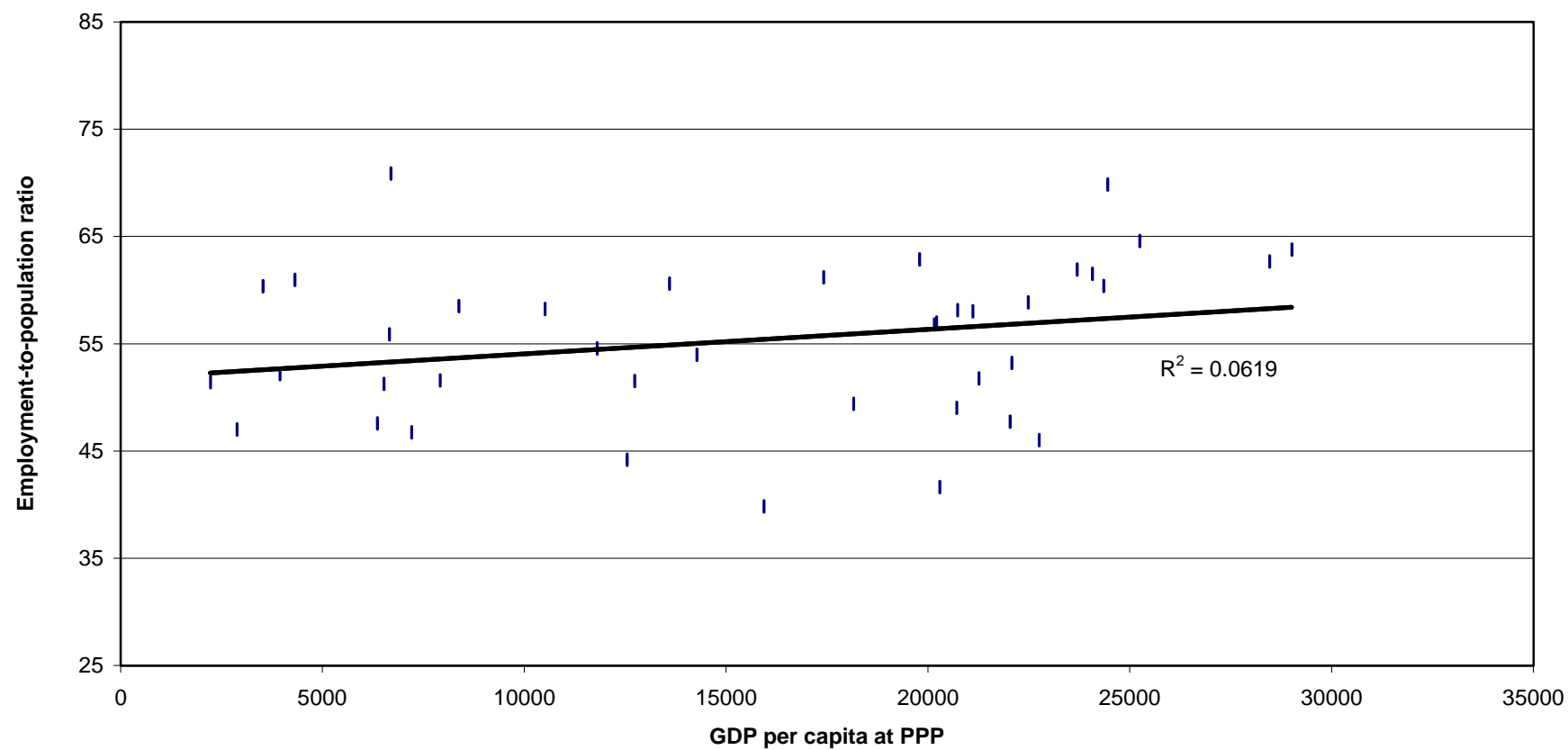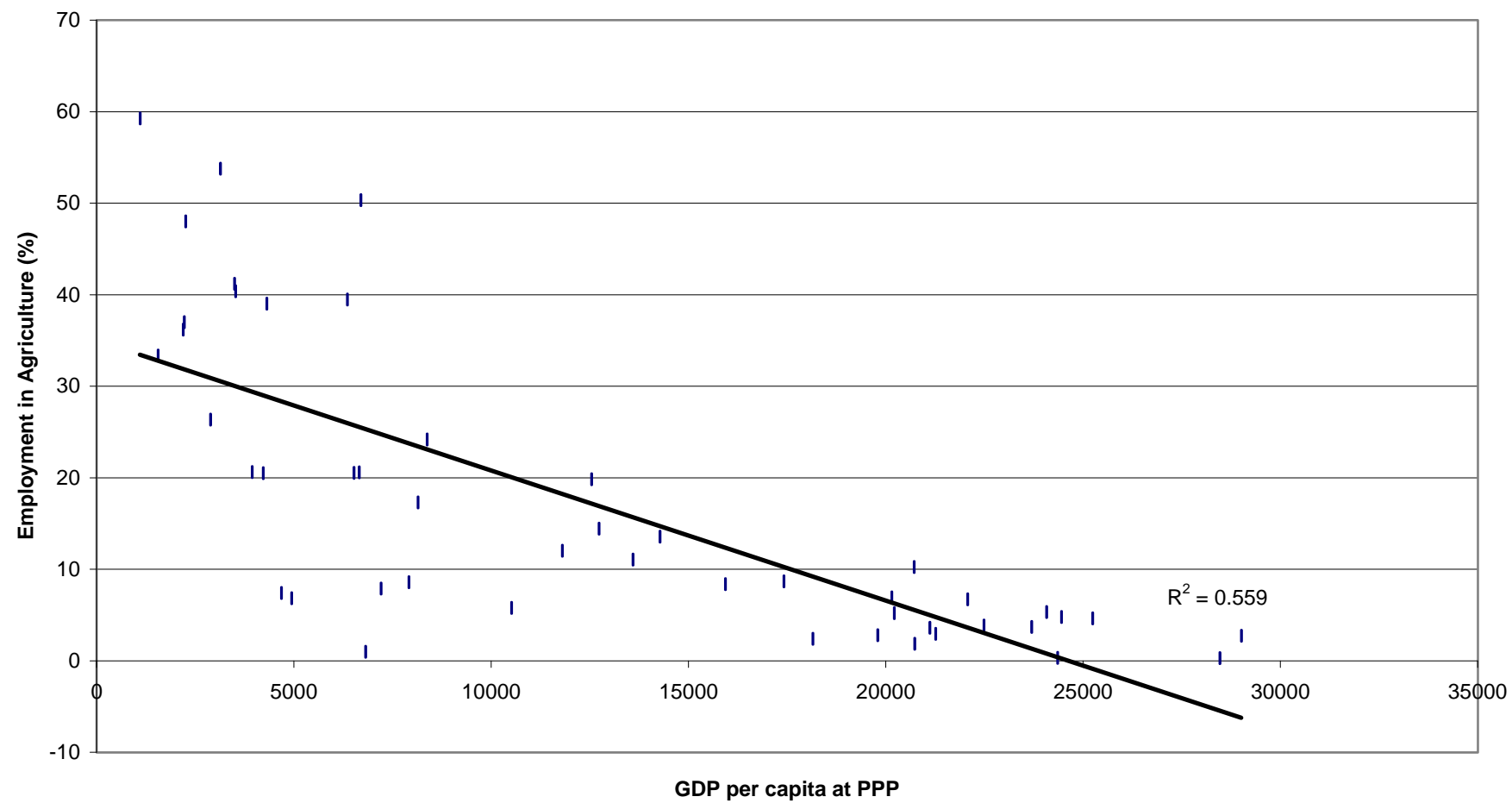3) KILM 4: Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), latest year (agriculture)
4) KILM 4: Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), latest year (services)
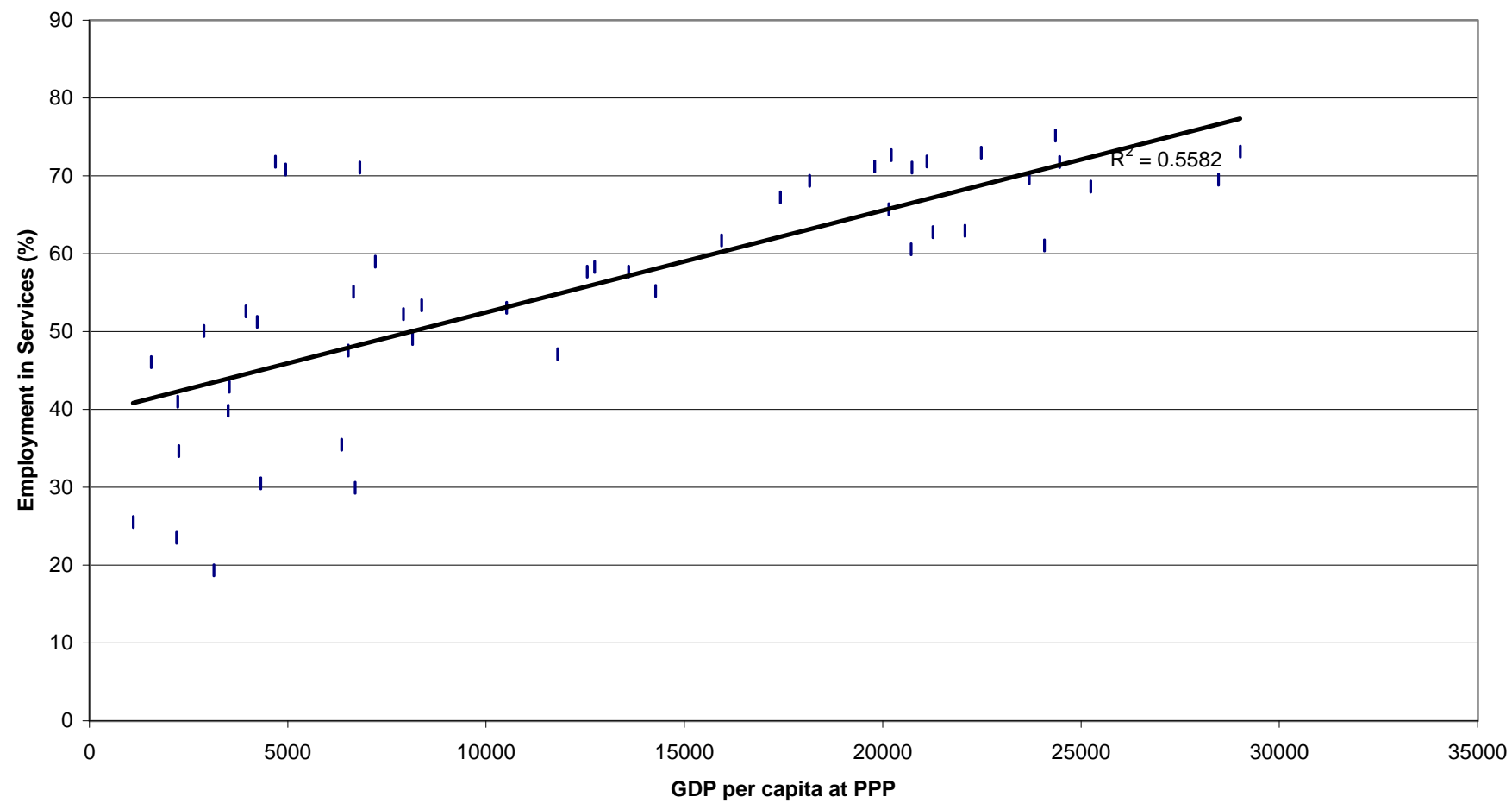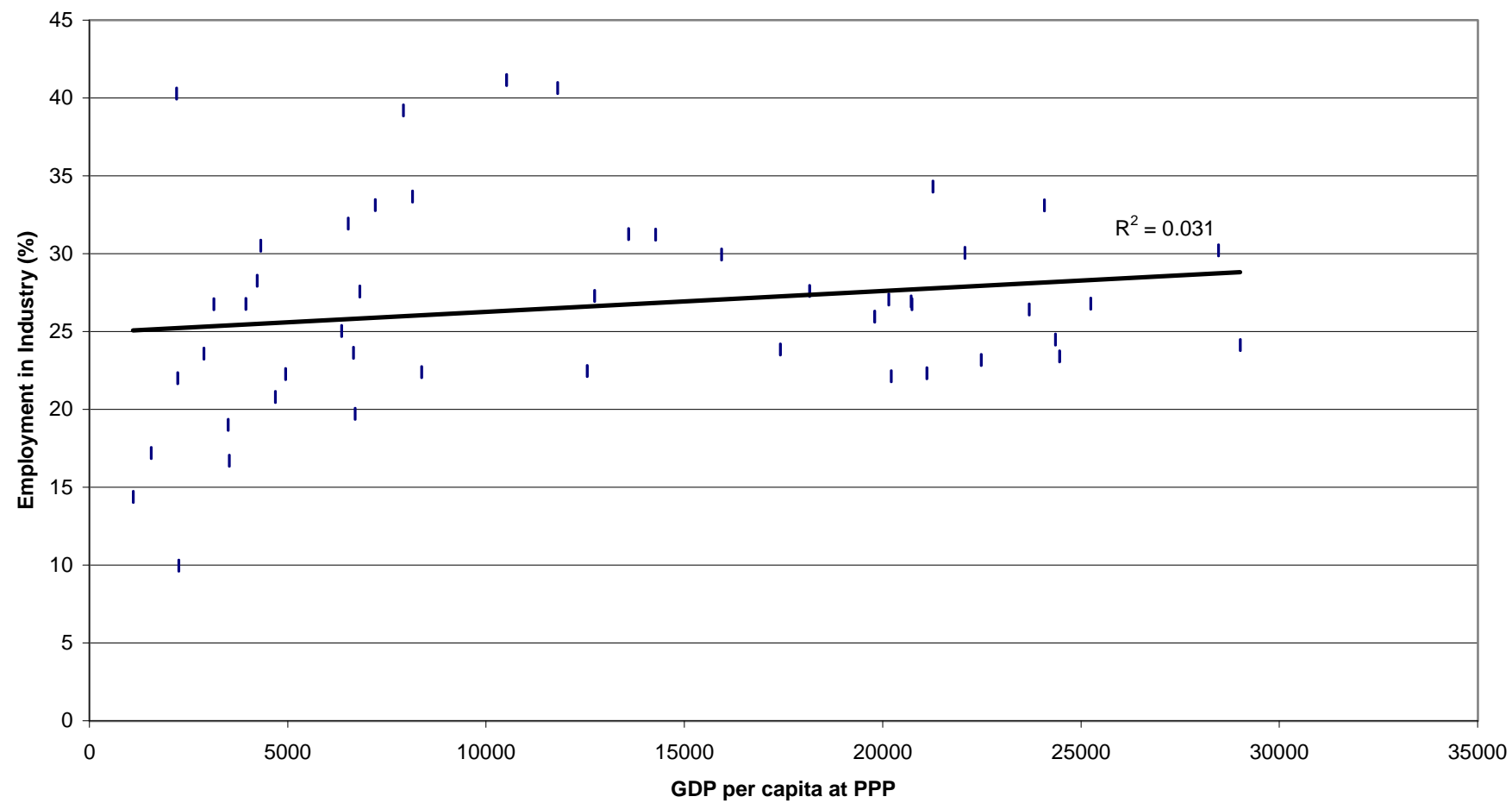5) KILM 4: Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), latest year (industry)
6) KILM 8: Unemployment rate of both sexes and GDP per capita at purchasing power parity (PPP), latest year
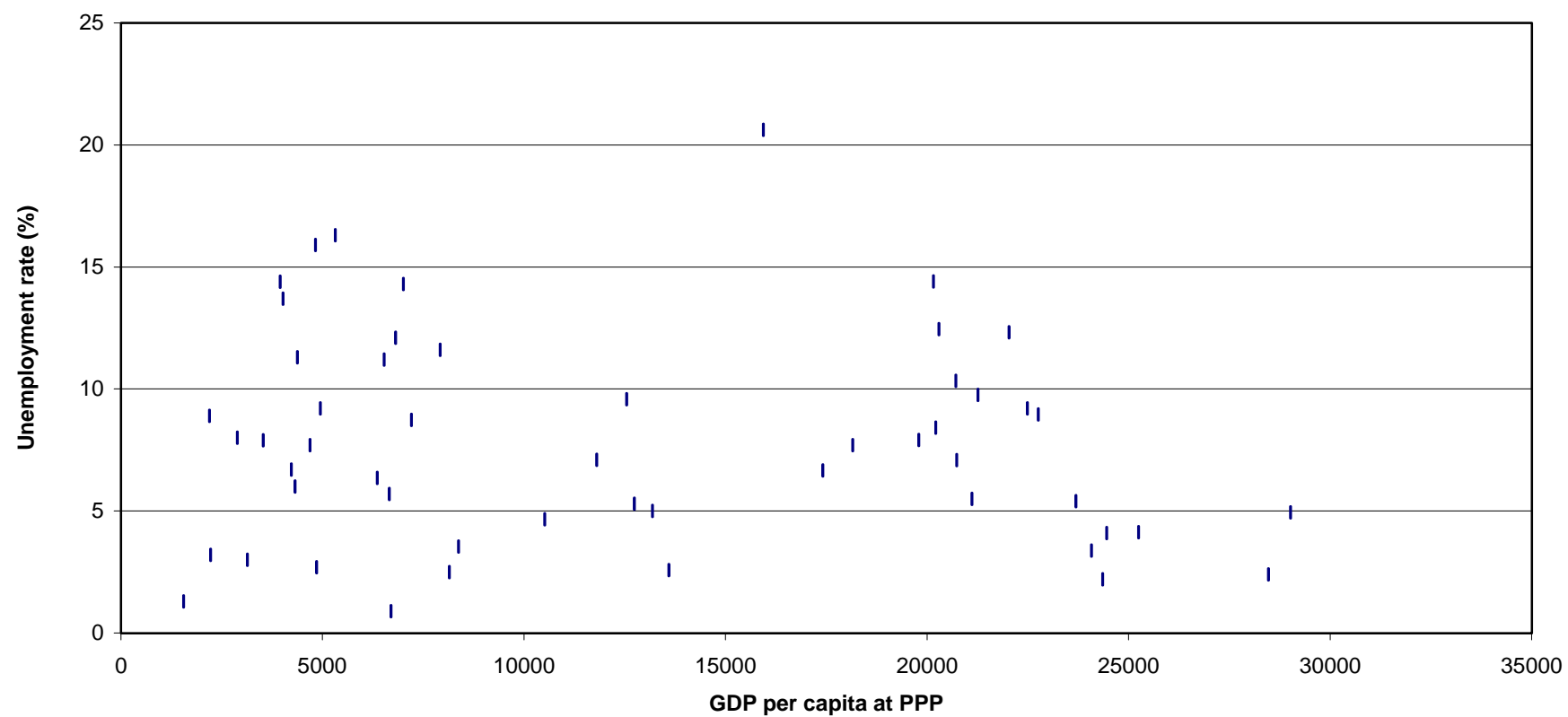
Figures - Regions

1) KILM 1: Labour force participation rate of both sexes aged 15 years and over, and GDP per capita at purchasing power parity (PPP), Developed (industrialized) countries, latest year
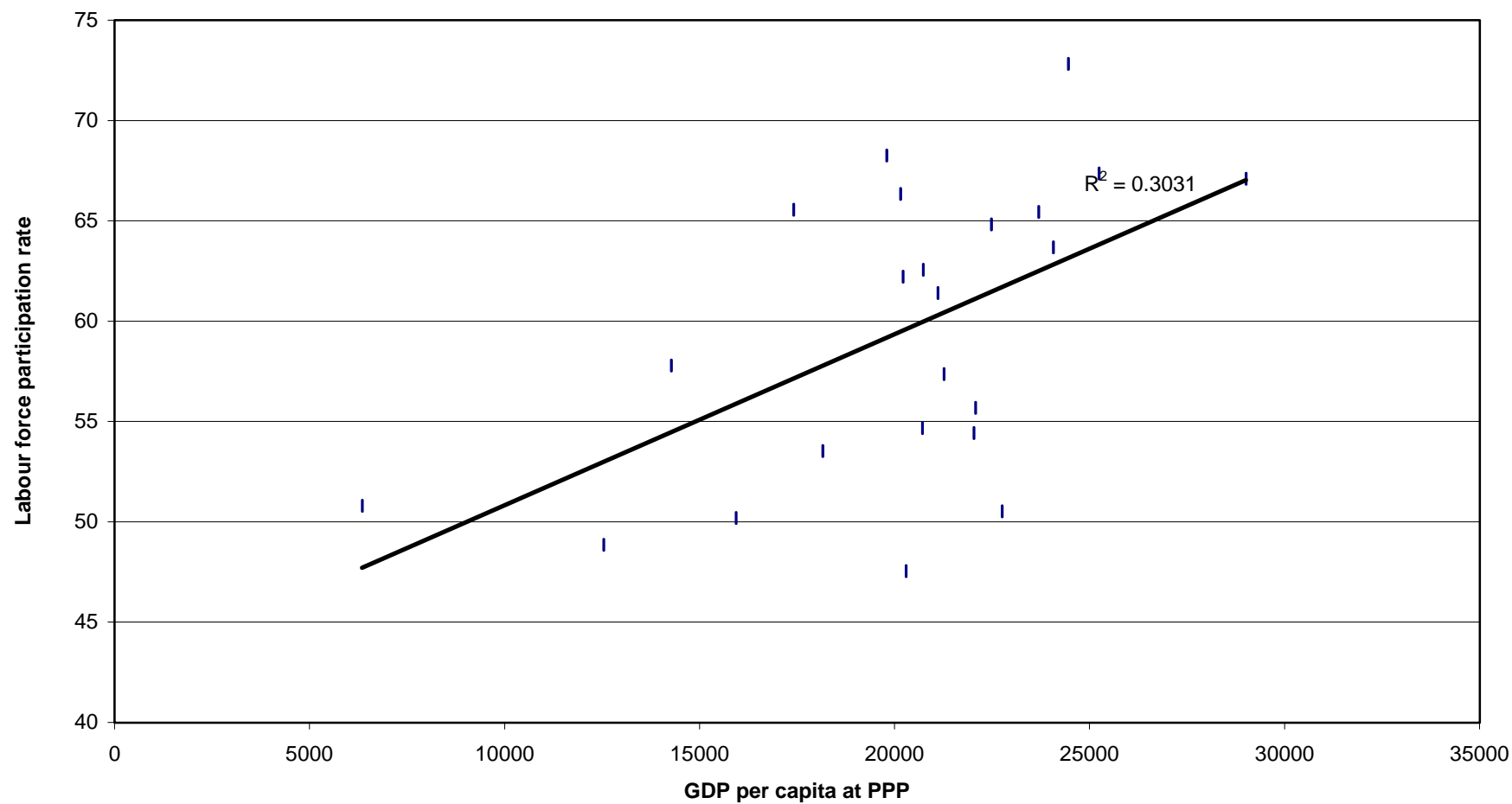2) KILM 1: Labour force participation rate of both sexes aged 15 years and over, and GDP per capita at purchasing power parity (PPP), Transition economies, latest year
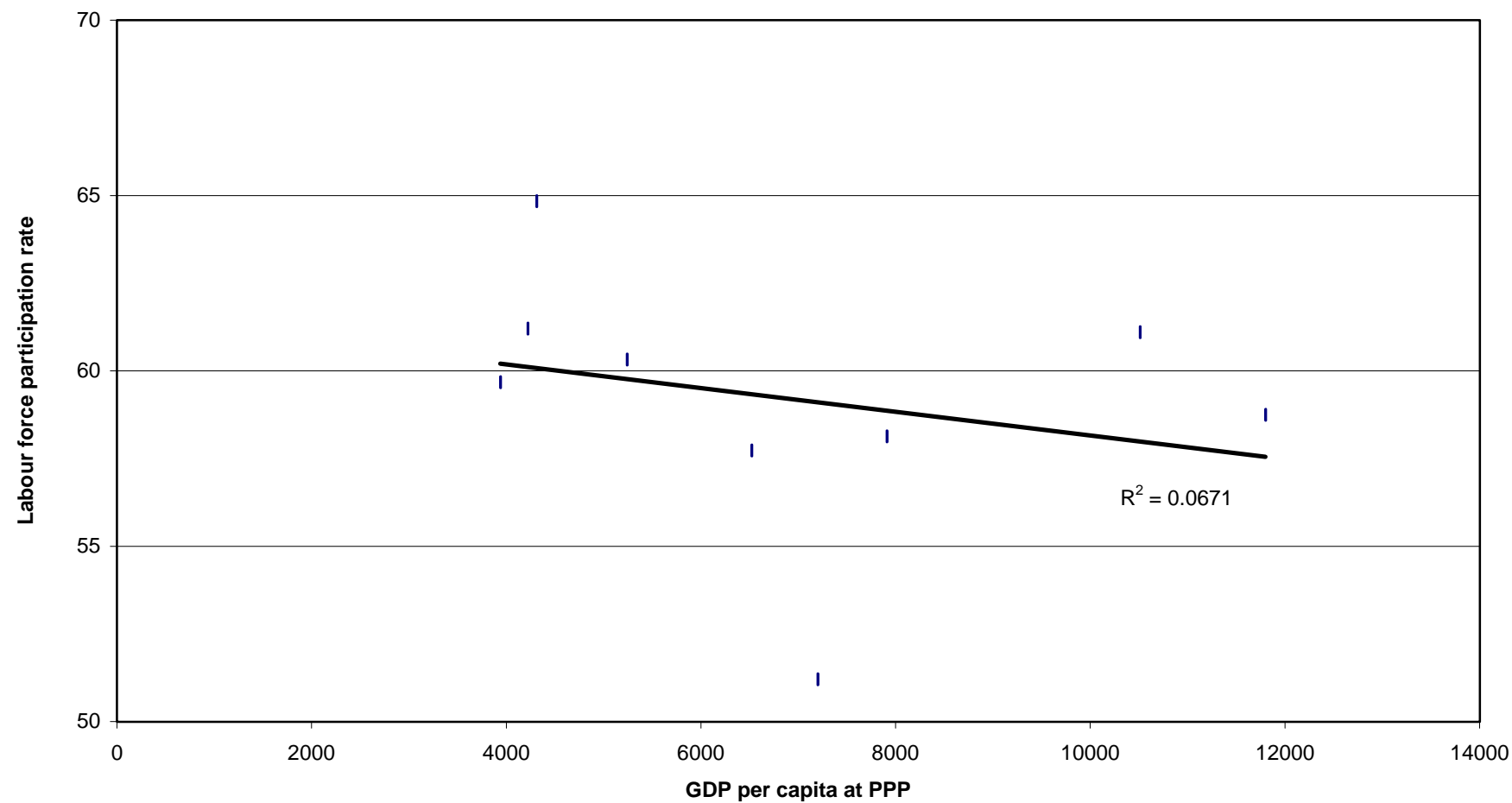3) KILM 1: Labour force participation rate of both sexes aged 15 years and over, and GDP per capita at purchasing power parity (PPP), Asia and the Pacific, latest year
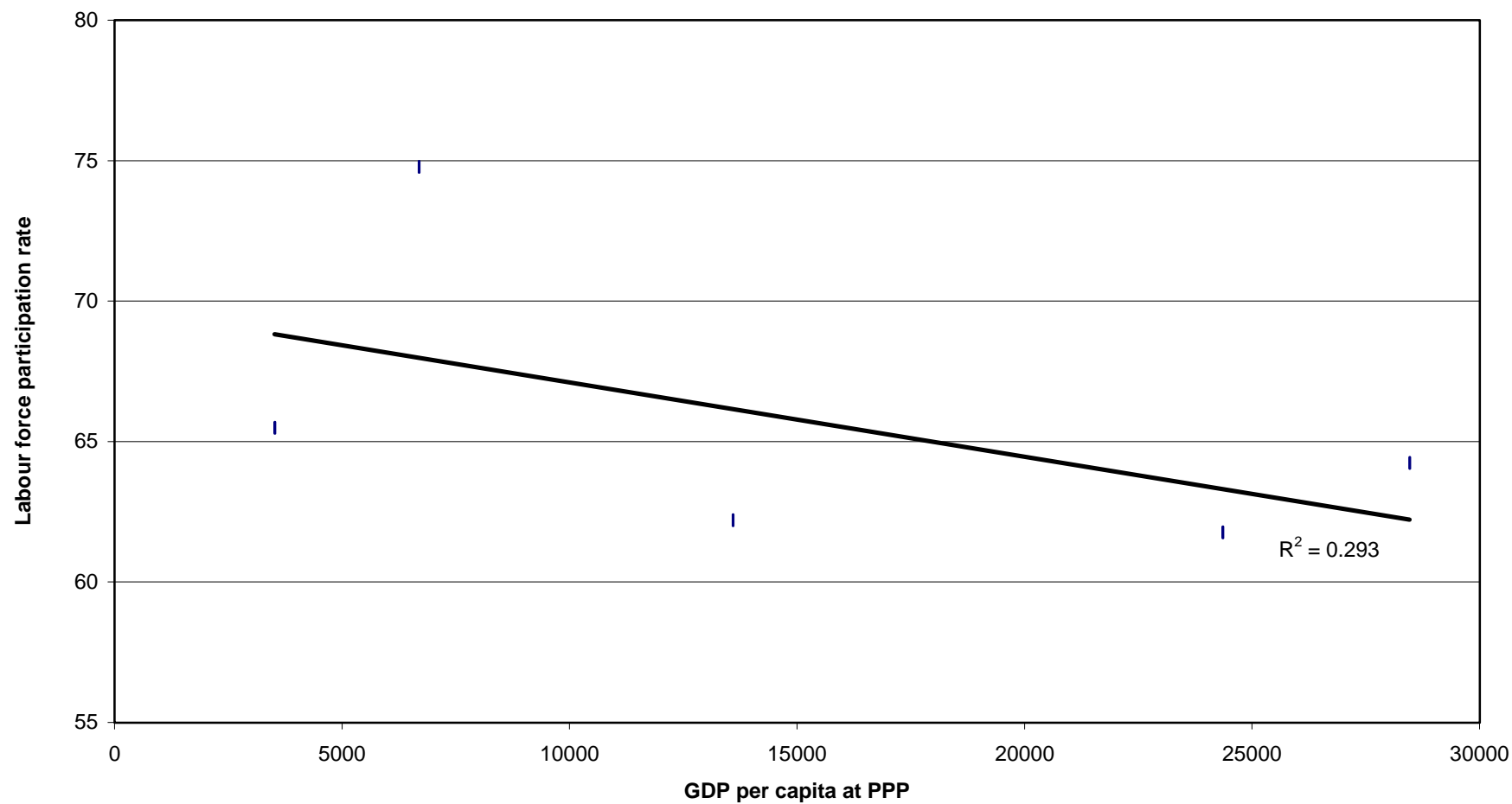4) KILM 1: Labour force participation rate of both sexes aged 15 years and over, and GDP per capita at purchasing power parity (PPP), Latin America and the Caribbean, latest year
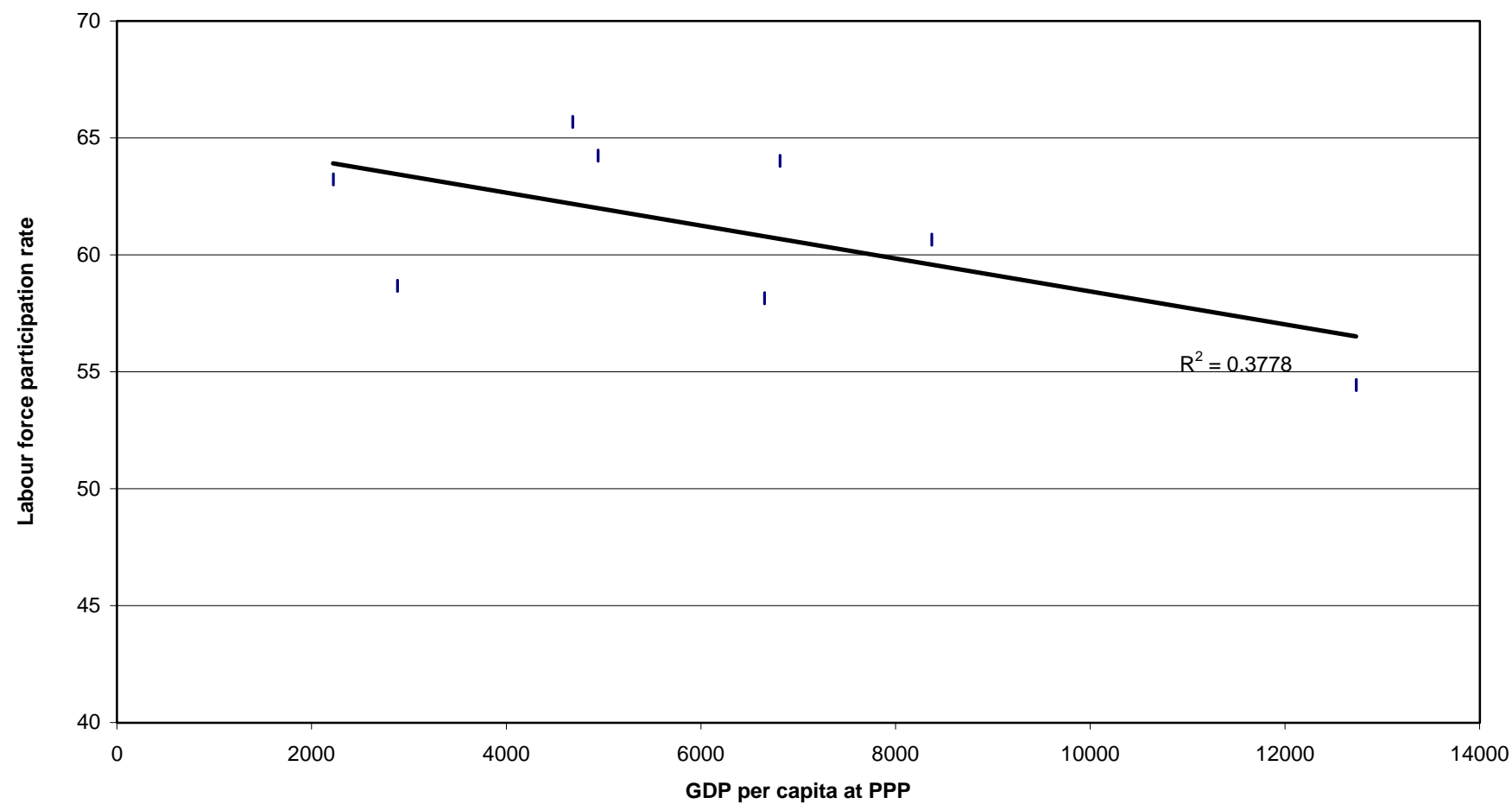5) KILM 2: Employment-to-population ratio of both sexes and GDP per capita at purchasing power parity (PPP), Developed (industrialized) countries, latest year
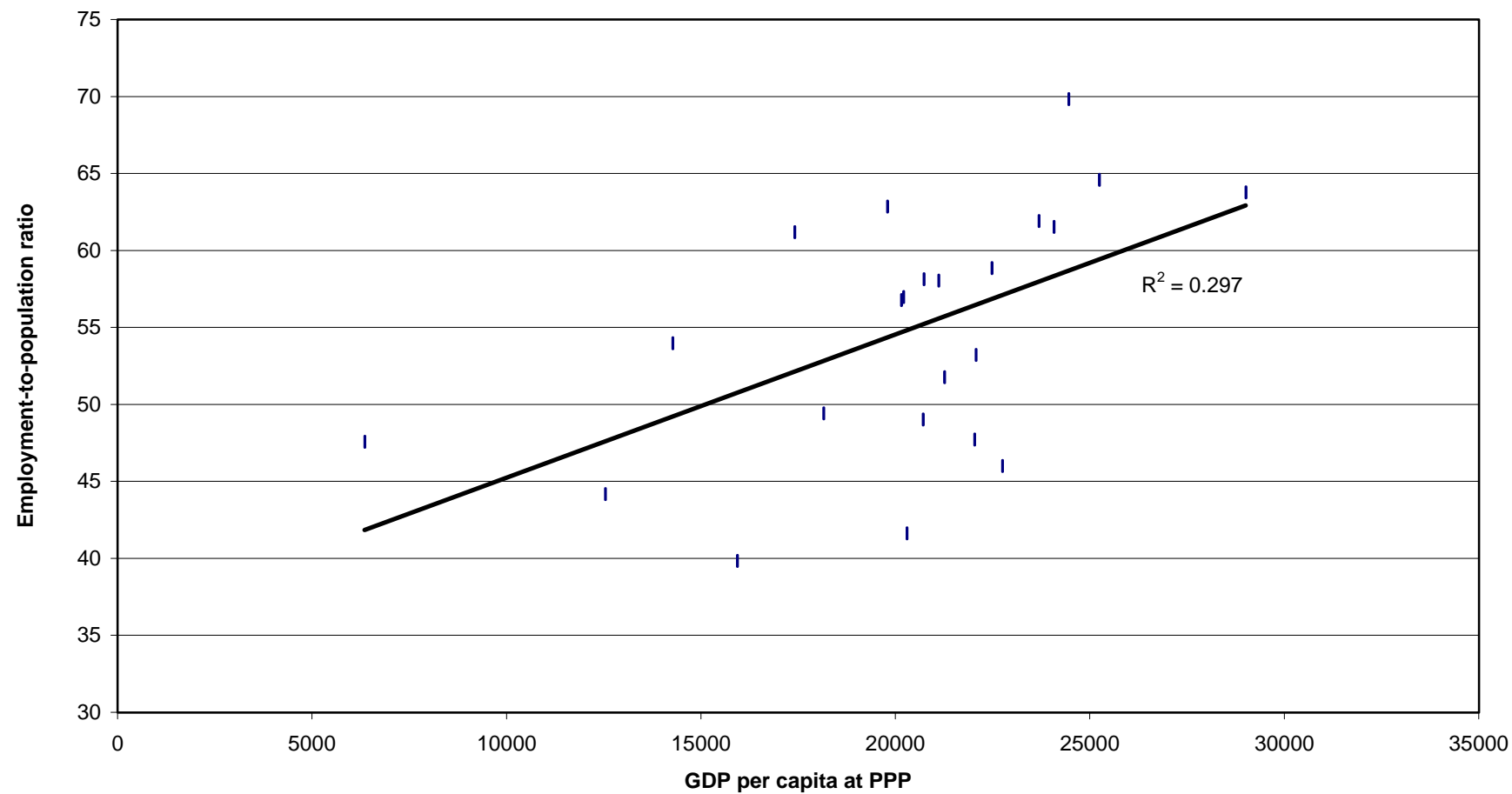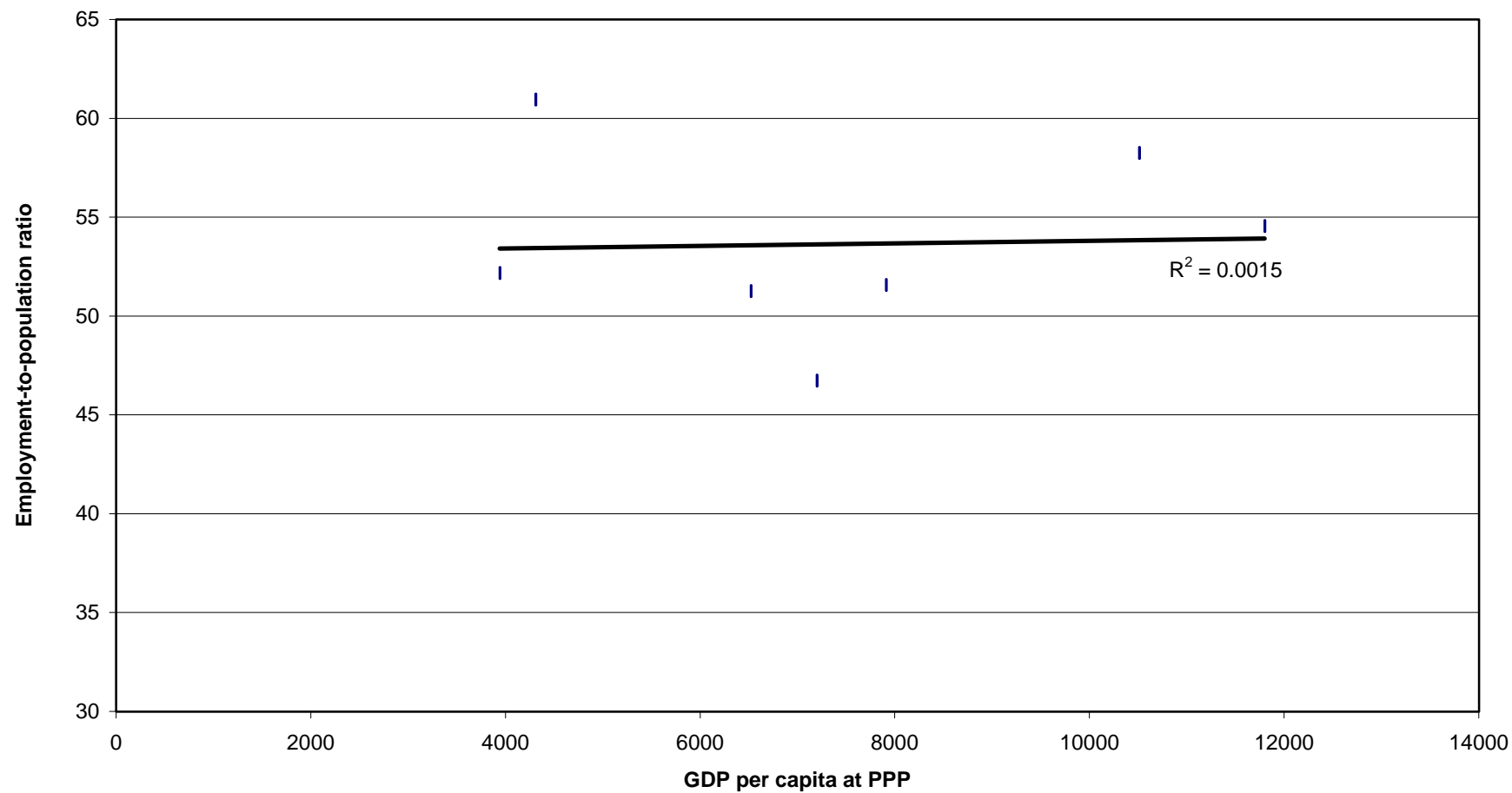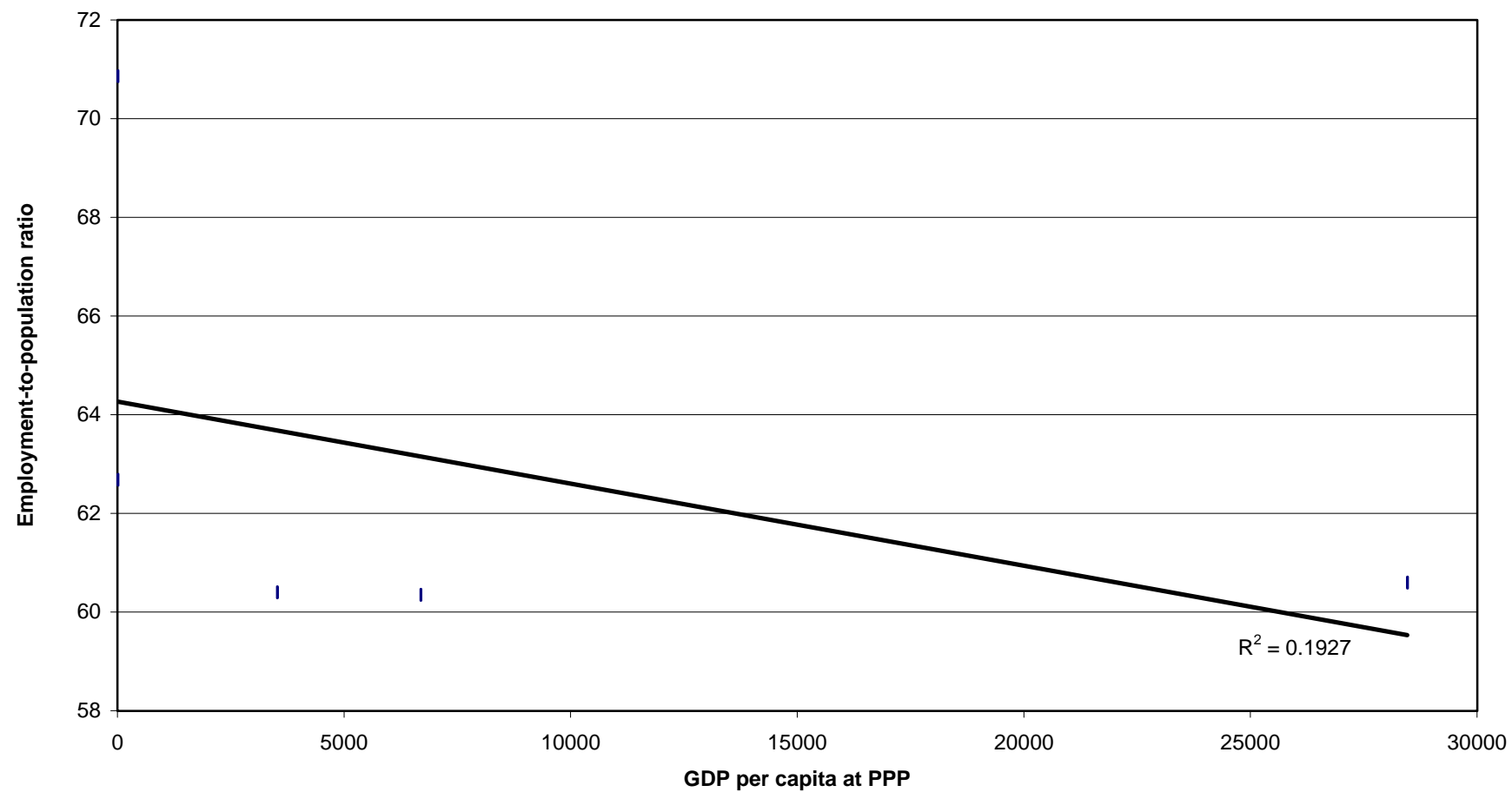6) KILM 2: Employment-to-population ratio of both sexes and GDP per capita at purchasing power parity (PPP), Transition economies, latest year
7) KILM 2: Employment-to-population ratio of both sexes and GDP per capita at purchasing power parity (PPP), Asia and the Pacific, latest year
8) KILM 2: Employment-to-population ratio of both sexes and GDP per capita at purchasing power parity (PPP), Latin America and the Caribbean, latest year
9) KILM 4 (agriculture): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Developed (industrialized) countries, latest year
10) KILM 4 (agriculture): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Transition economies, latest year
11) KILM 4 (agriculture): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Asia and the Pacific, latest year
12) KILM 4 (agriculture): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Latin America and the Caribbean, latest year
13) KILM 4 (services): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Developed (industrialized) countries, latest year
14) KILM 4 (services): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Transition economies, latest year
15) KILM 4 (services): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Asia and the Pacific, latest year
16) KILM 4 (services): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Latin America and the Caribbean, latest year
17) KILM 4 (industry): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Developed (industrialized) countries, latest year
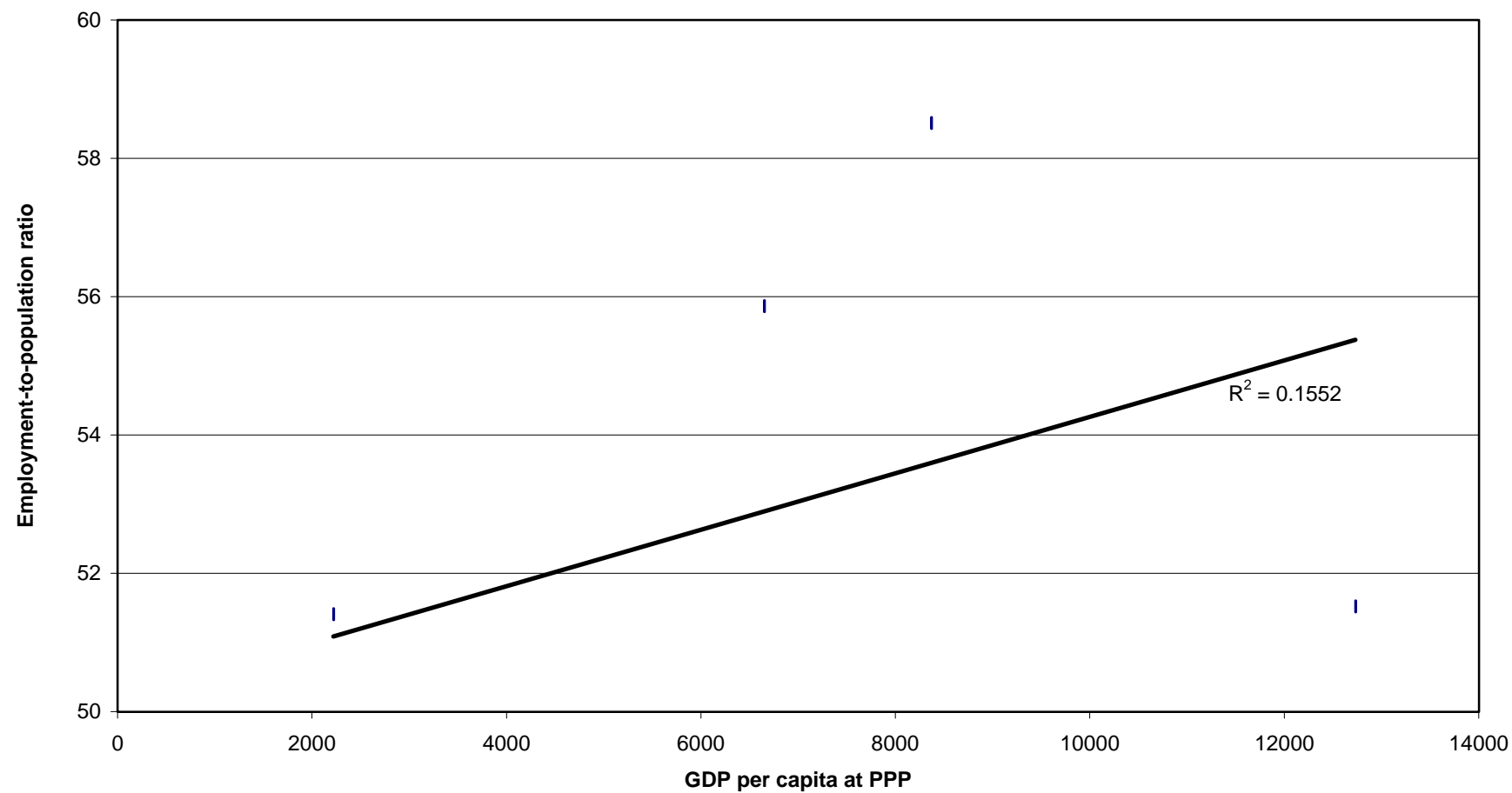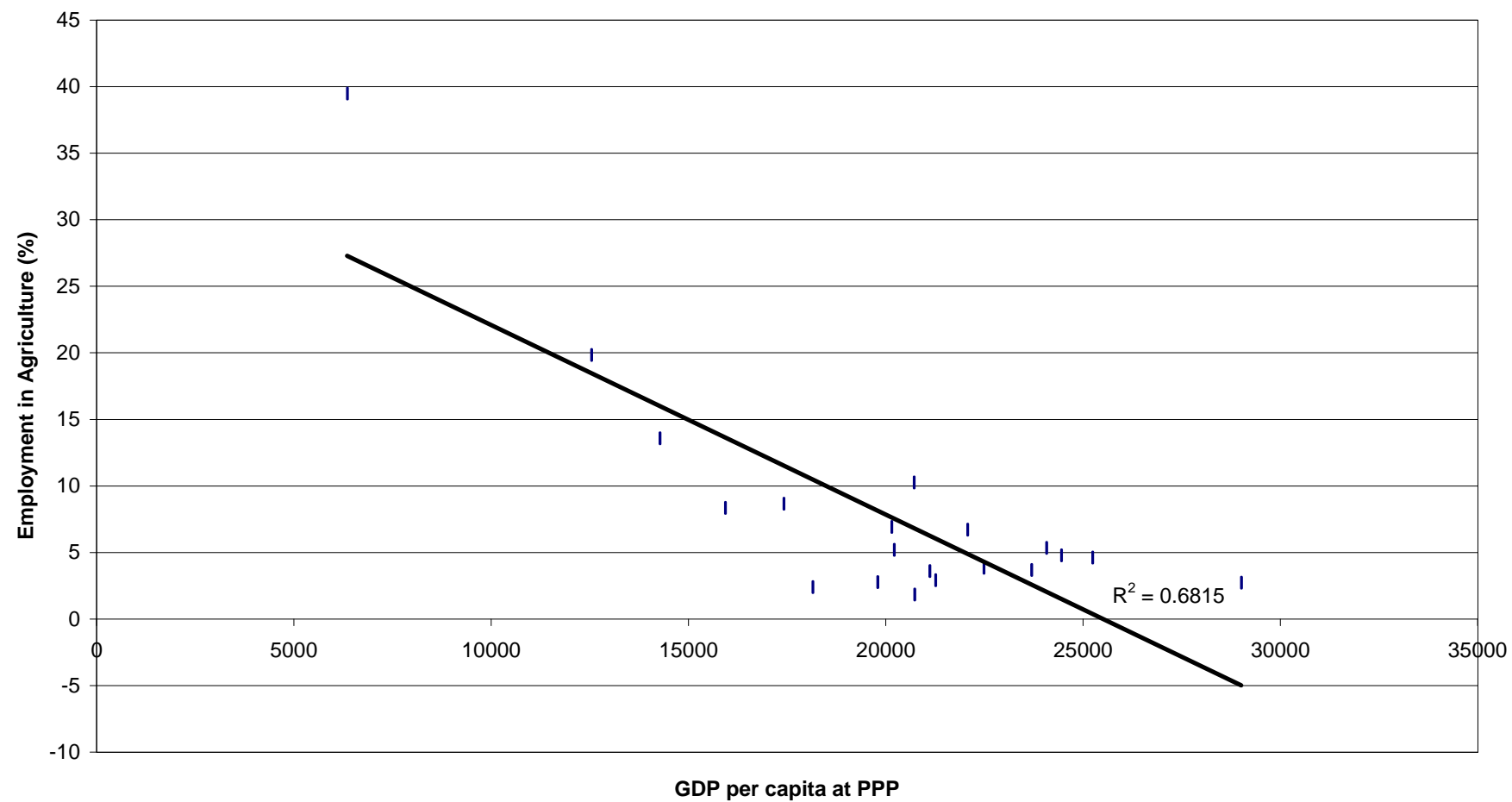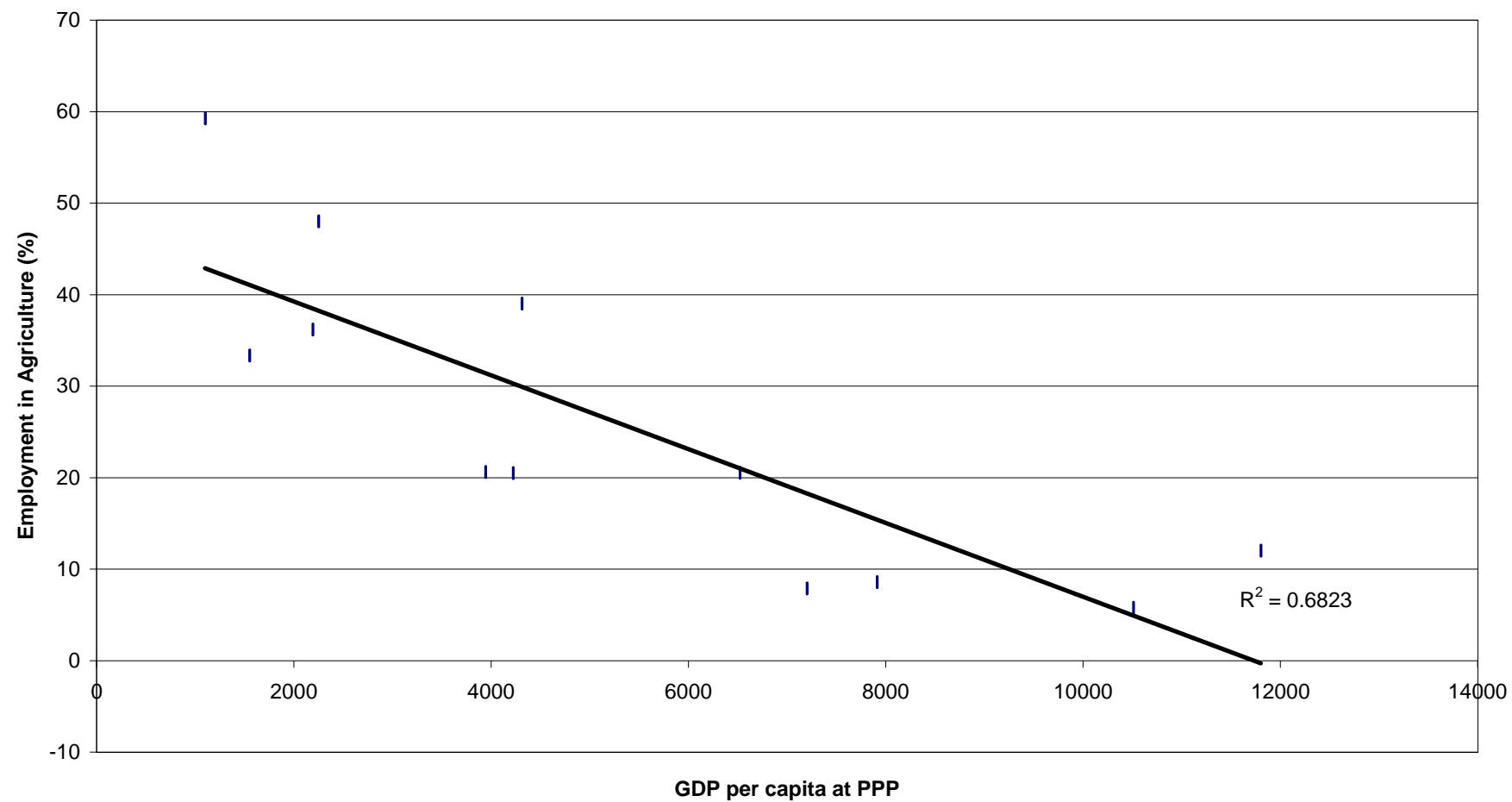18) KILM 4 (industry): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Transition economies, latest year
19) KILM 4 (industry): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Asia and the Pacific, latest year
20) KILM 4 (industry): Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), Latin America and the Caribbean, latest year
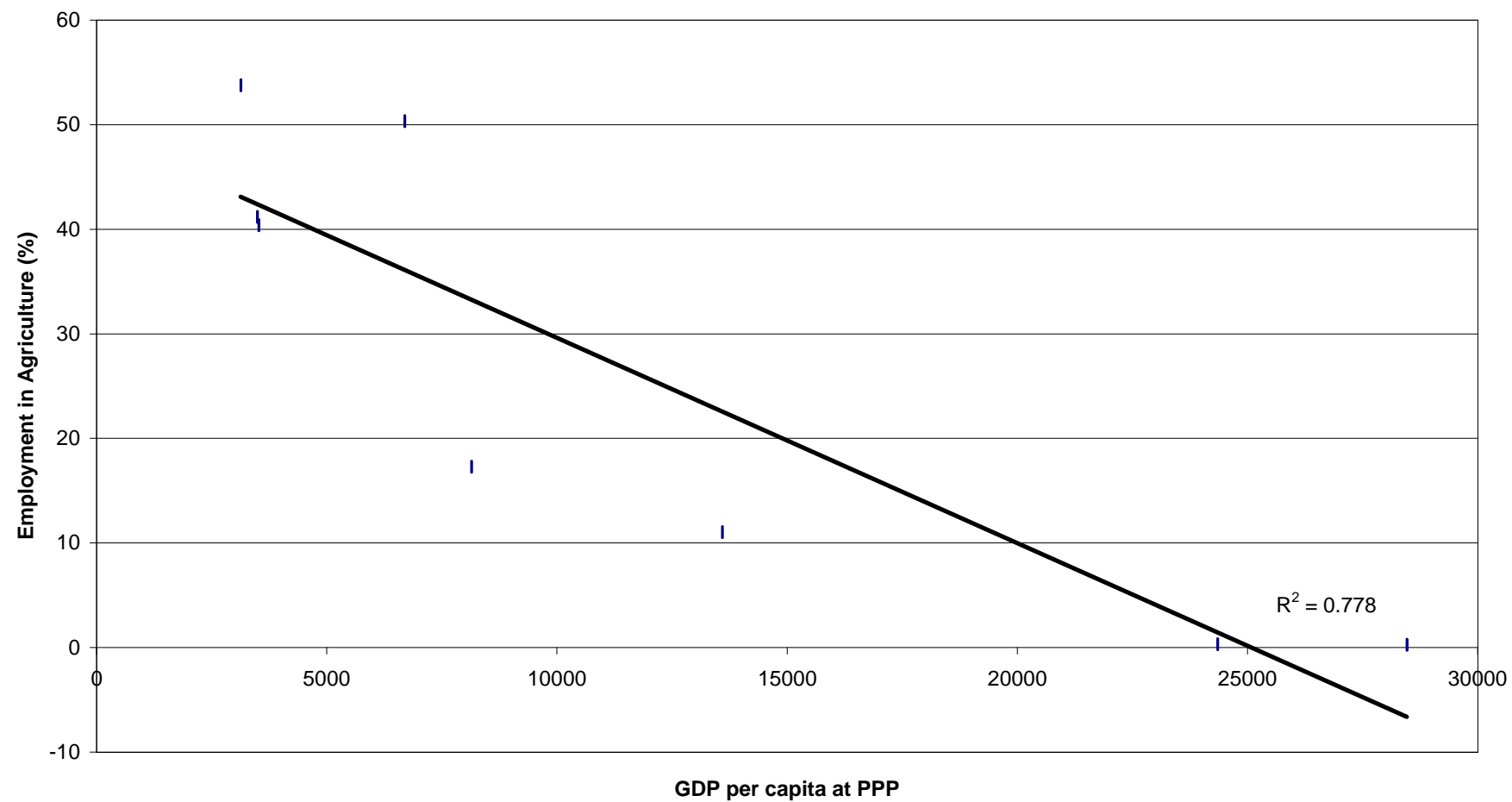21) KILM 8: Unemployment rate of both sexes and GDP per capita at purchasing power parity (PPP), Developed (industrialized) countries, latest year

Figures – Gender

Figure 1. Labour force participation rate of both sexes aged 15 years and over, and GDP per capita at purchasing power parity (PPP), latest year

$R^2 = 0.0022$

Labour force participation rate

GDP per capita at PPP

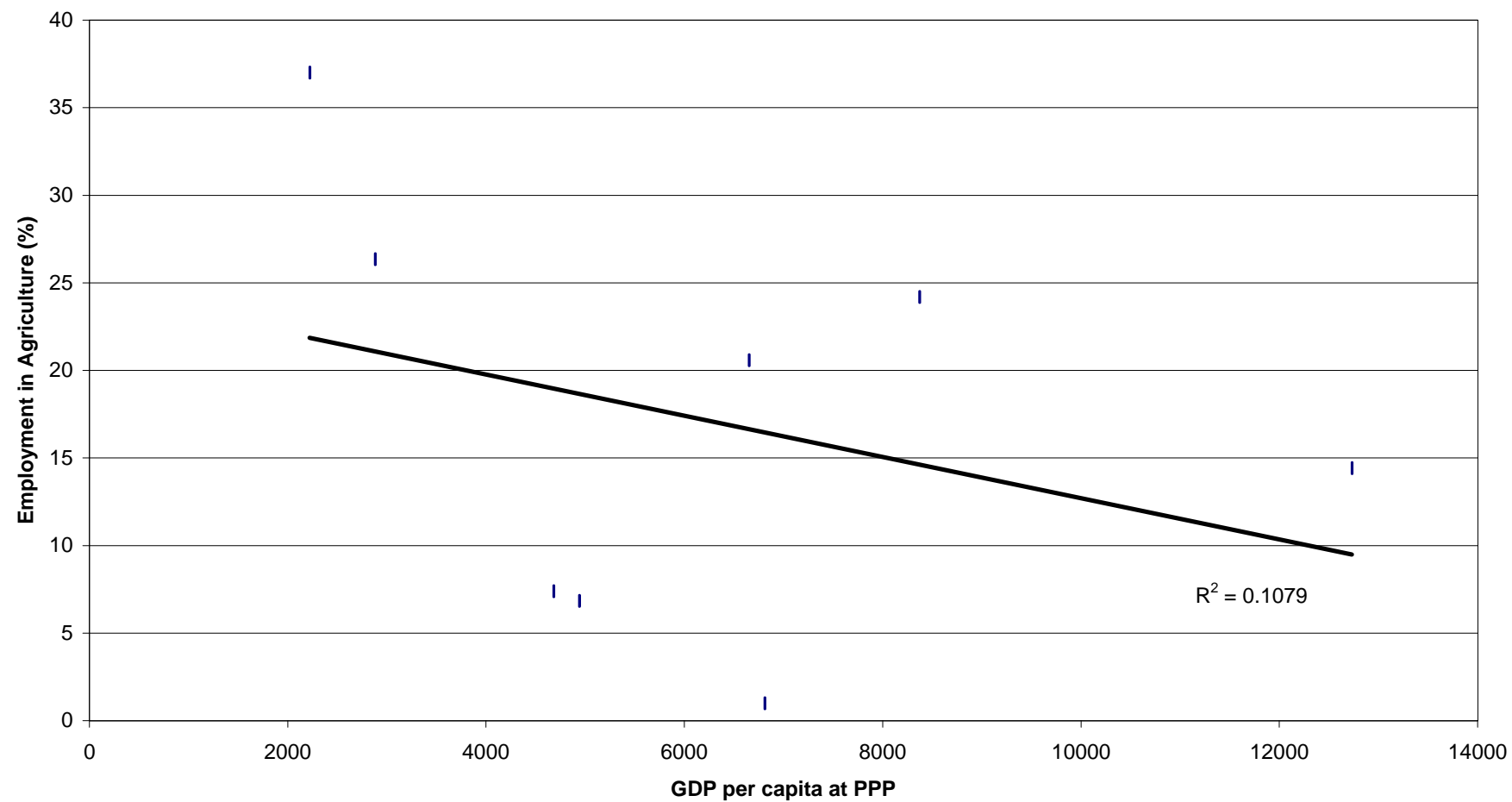Figure 2. Employment-to-population ratio of both sexes and GDP per capita at purchasing power parity (PPP), latest year

Figure 3. Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), latest year

Figure 4. Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), latest year

$R^2 = 0.5582$

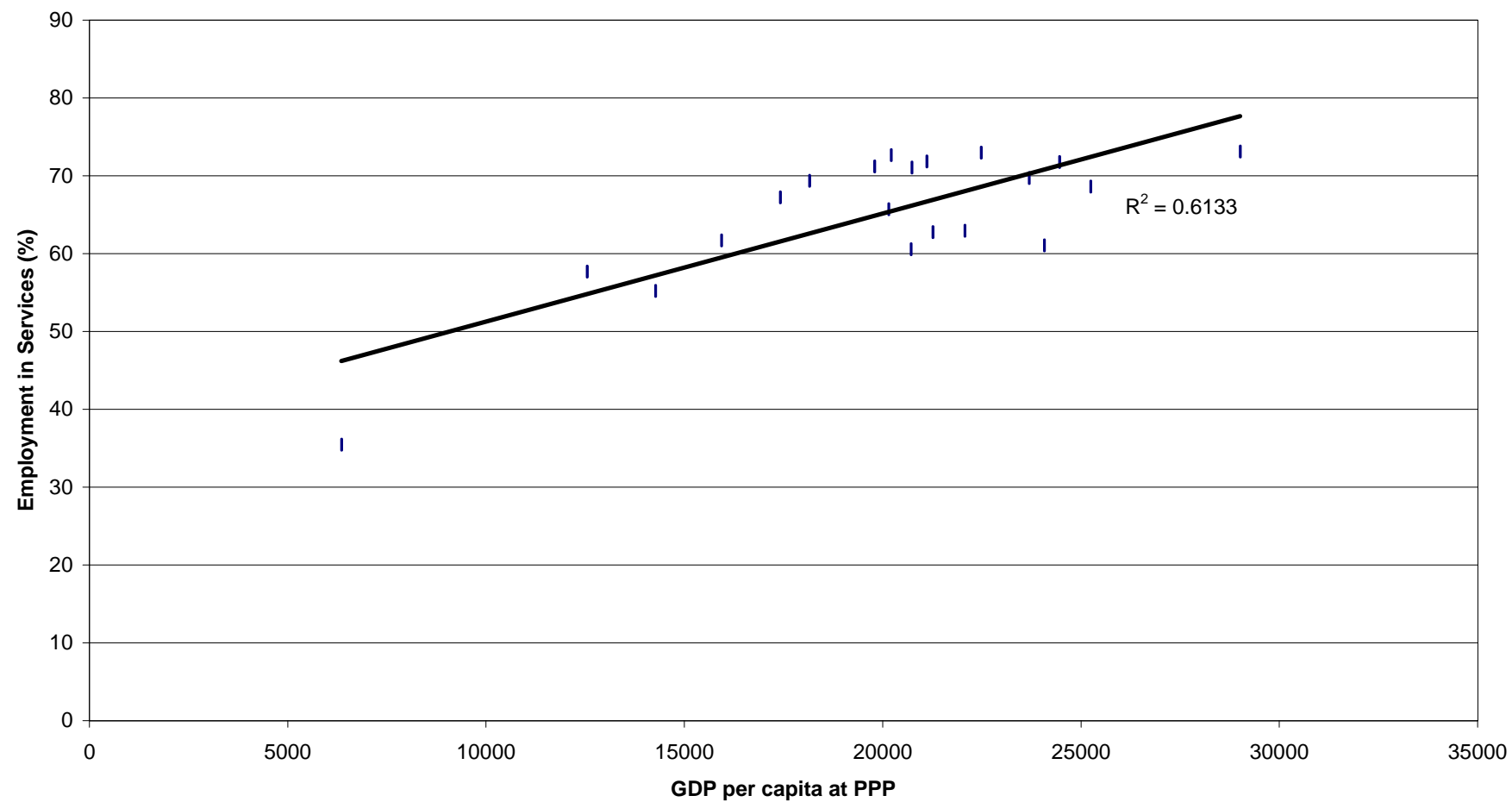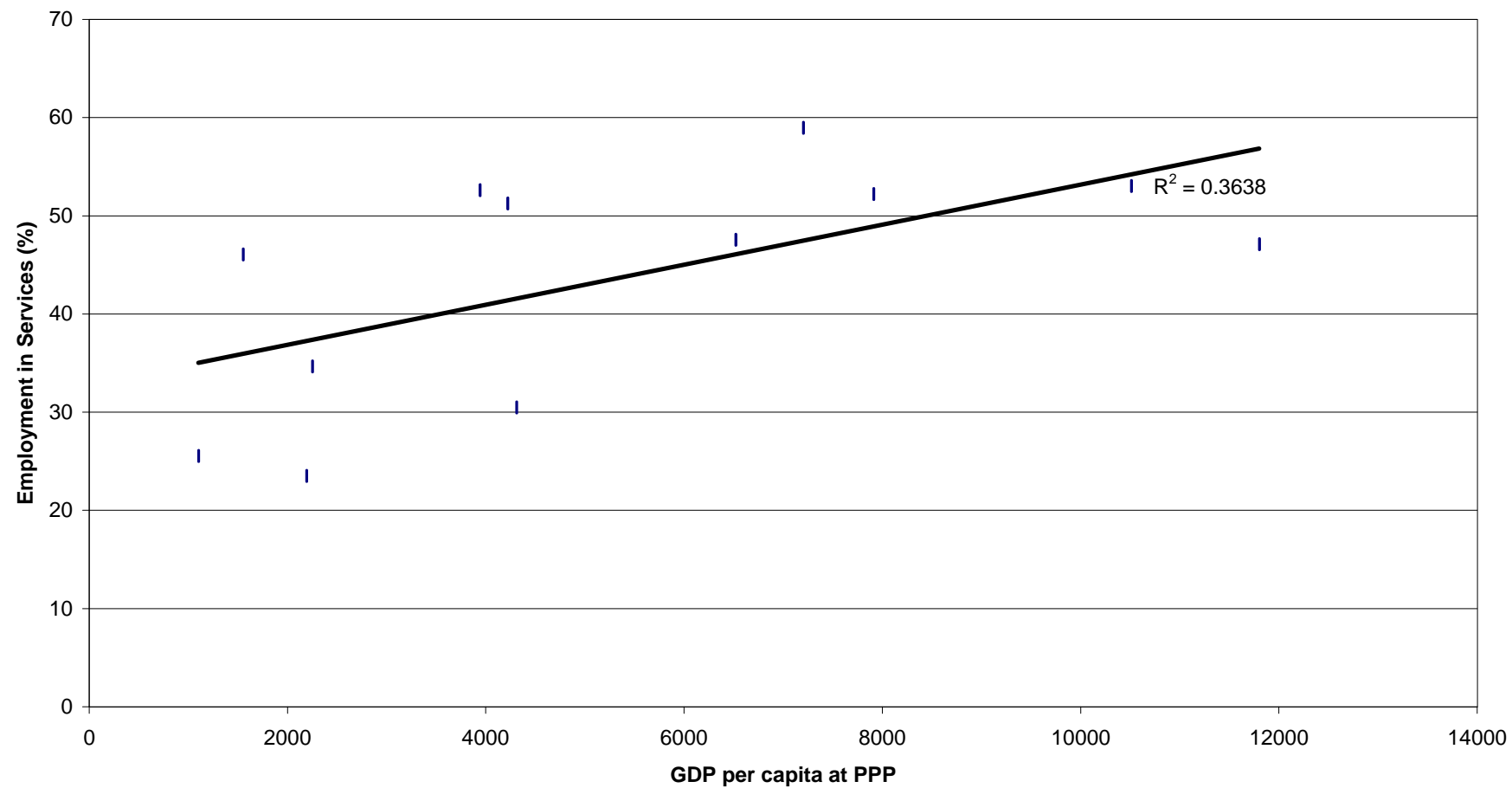**Figure 5. Employment by sector of both sexes and GDP per capita at purchasing power parity (PPP), latest year**

$R^2 = 0.031$

Employment in Industry (%)

GDP per capita at PPP

Figure 6. Unemployment rate of both sexes and GDP per capita at purchasing power parity (PPP), latest year

Figure 7. Labour force participation rate of both sexes aged 15 years and over, and GDP per capita at purchasing power parity (PPP), Developed (Industrialized) countries, latest year

$R^2 = 0.3031$

GDP per capita at PPP

Labour force participation rate

Figure 8. Labour force participation rate of both sexes aged 15 years and over, and GDP per capita at purchasing power parity (PPP), Transition Economies, latest year

$R^2 = 0.0671$

Labour force participation rate

GDP per capita at PPP

Figure 9. Labour force participation rate of both sexes aged 15 years and over, and GDP per capita at purchasing power parity (PPP), Asia and the Pacific, latest year

$R^2 = 0.293$

Labour force participation rate

GDP per capita at PPP

Figure10. Labour force participation rate of both sexes aged 15 years and over, and GDP per capita at purchasing power parity (PPP), Latin America and the Caribbean, latest year

$R^2 = 0.3778$

Labour force participation rate

GDP per capita at PPP

Figure 11. Employment-to-population ratio of both sexes, and GDP per capita at purchasing power parity (PPP), Developed (Industrialized) countries, latest year
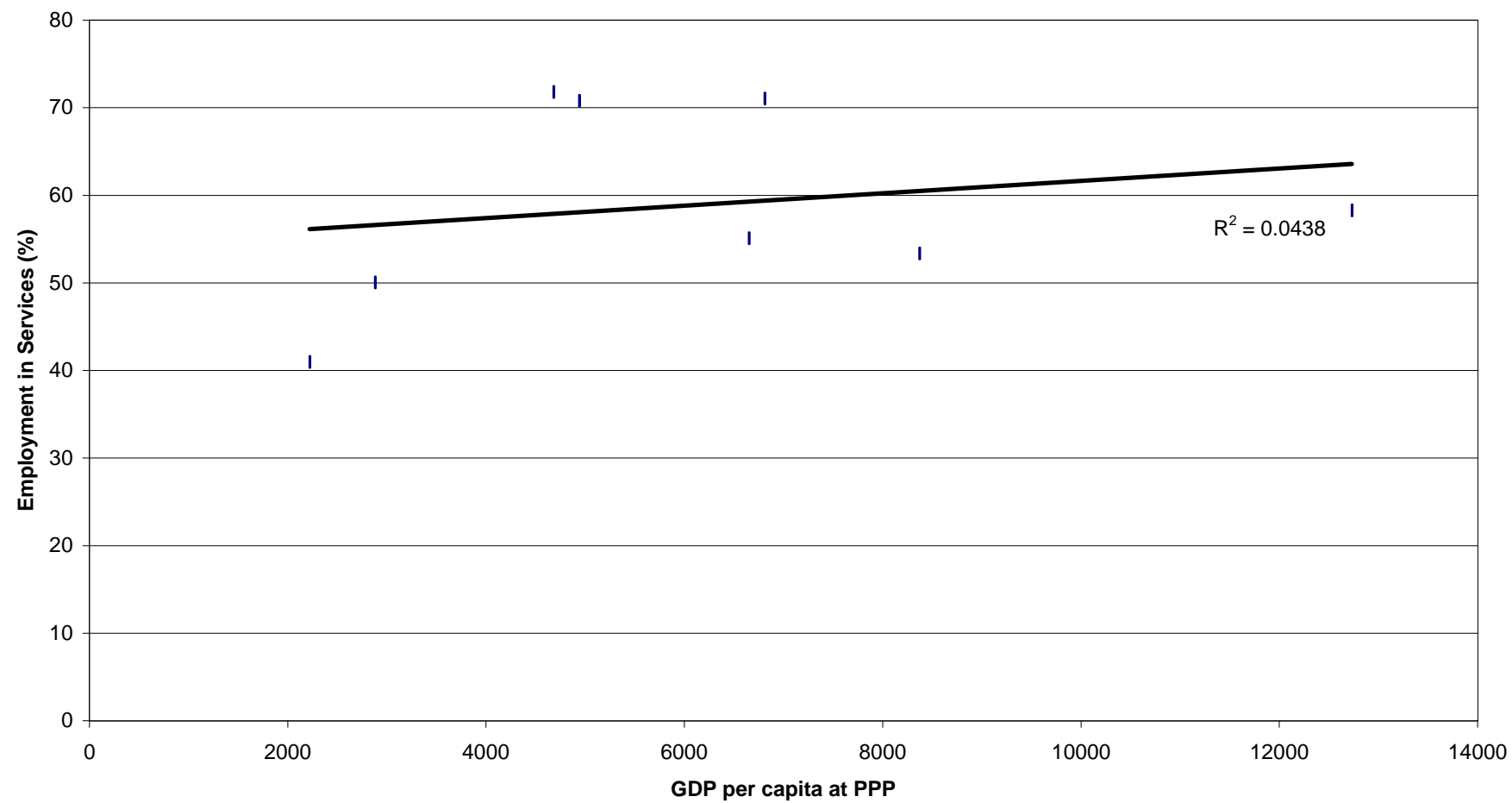
Figure 12. Employment-to-population ratio of both sexes, and GDP per capita at purchasing power parity (PPP), Transition Economies, latest year
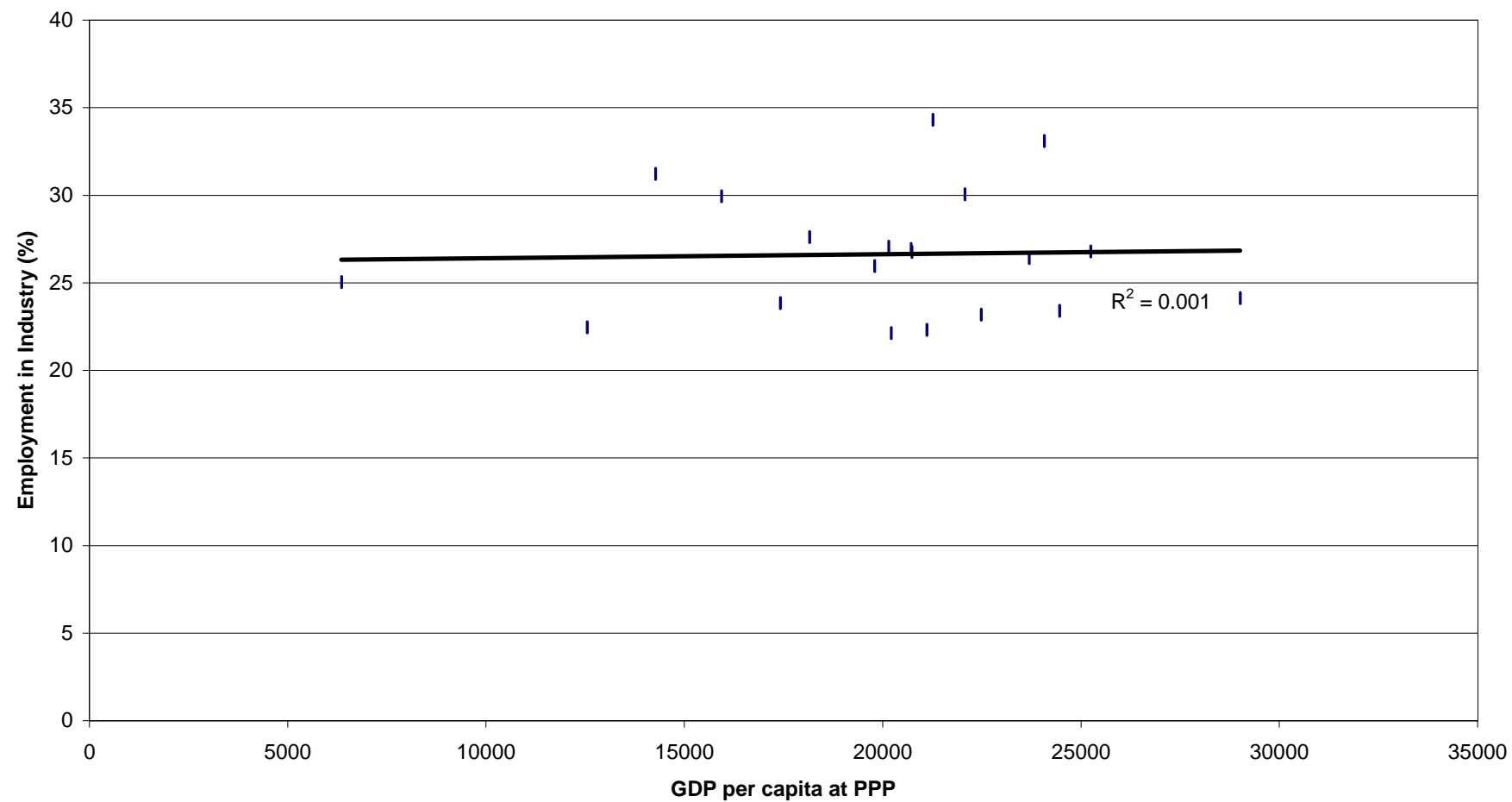
$R^2 = 0.0015$

Employment-to-population ratio

GDP per capita at PPP

Figure 13. Employment-to-population ratio of both sexes, and GDP per capita at purchasing power parity (PPP), Asia and the Pacific, latest year
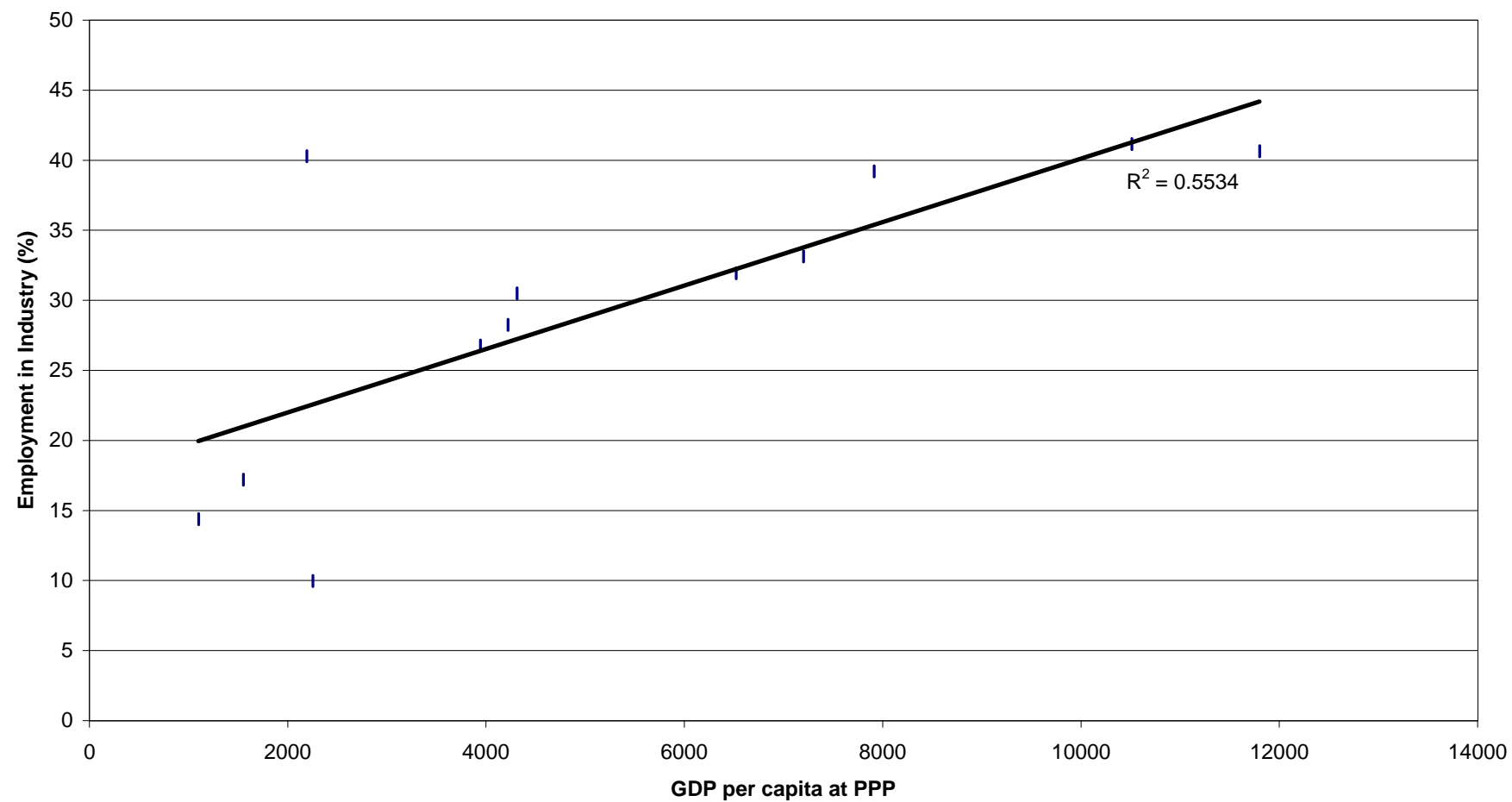
$R^2 = 0.1927$

Figure 14. Employment-to-population ratio of both sexes, and GDP per capita at purchasing power parity (PPP),
Latin America and the Caribbean, latest year

$R^2 = 0.1552$
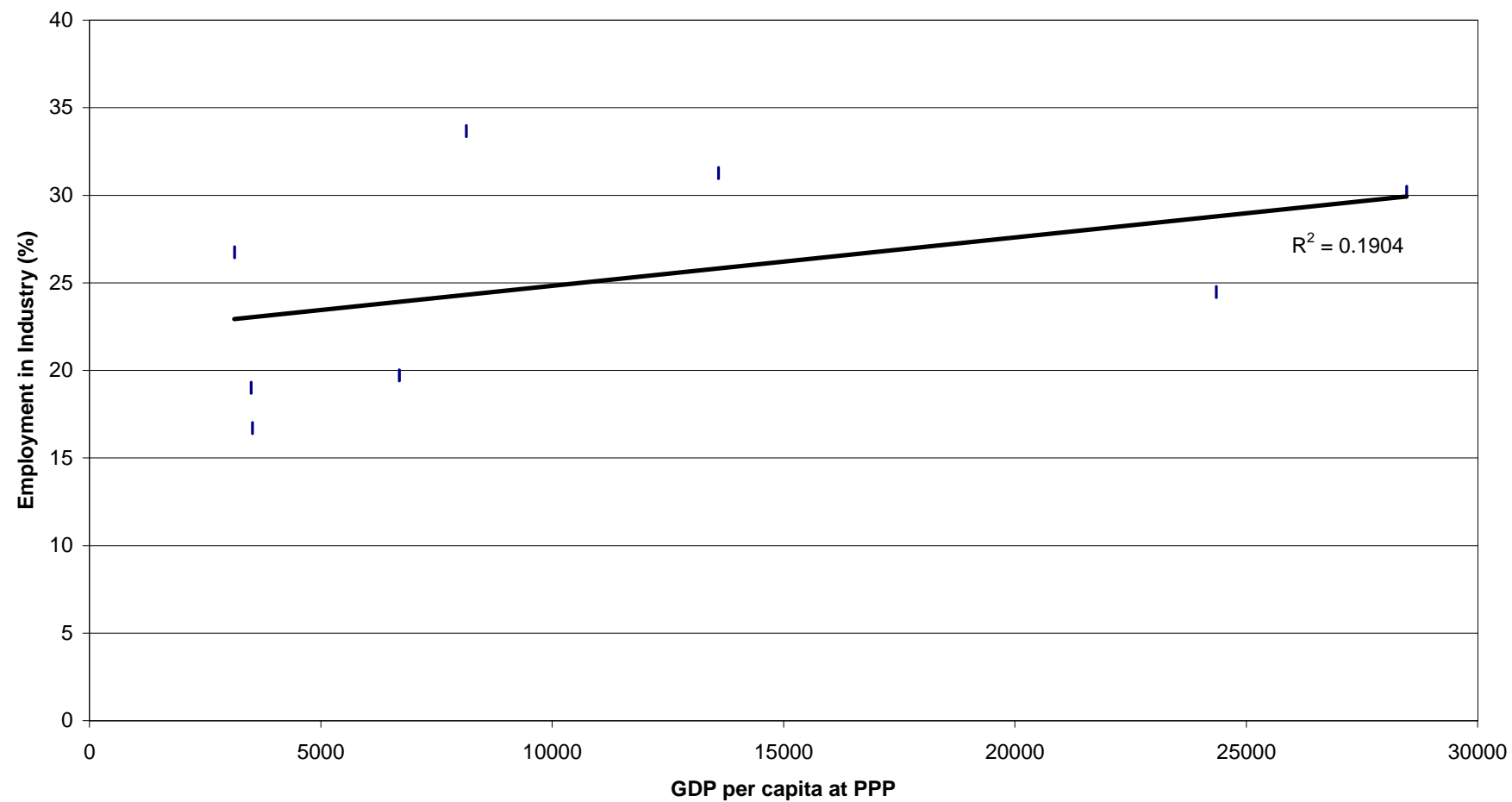
Employment-to-population ratio

GDP per capita at PPP

Figure 15. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Developed (Industrialized) countries, latest year
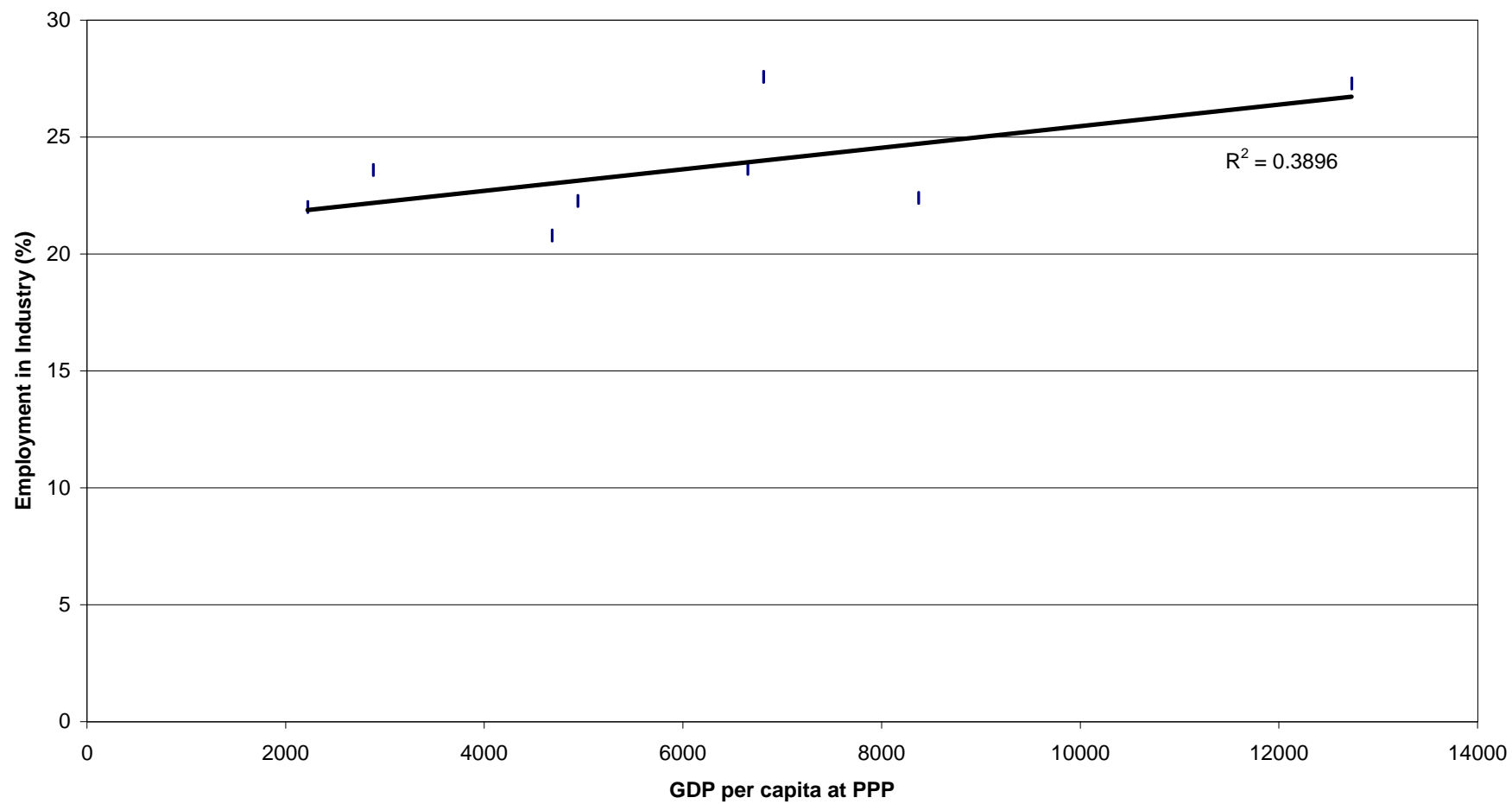
Employment in Agriculture (%)

GDP per capita at PPP

$R^2 = 0.6815$

Figure 16. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Transition Economies, latest year

$R^2 = 0.6823$

Employment in Agriculture (%)

GDP per capita at PPP

Figure 17. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Asia and the Pacific, latest year

$R^2 = 0.778$

Employment in Agriculture (%)

GDP per capita at PPP

Figure 18. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Latin America and the Caribbean, latest year

$R^2 = 0.1079$

Employment in Agriculture (%)

GDP per capita at PPP

Figure 19. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Developed (Industrialized) countries, latest year
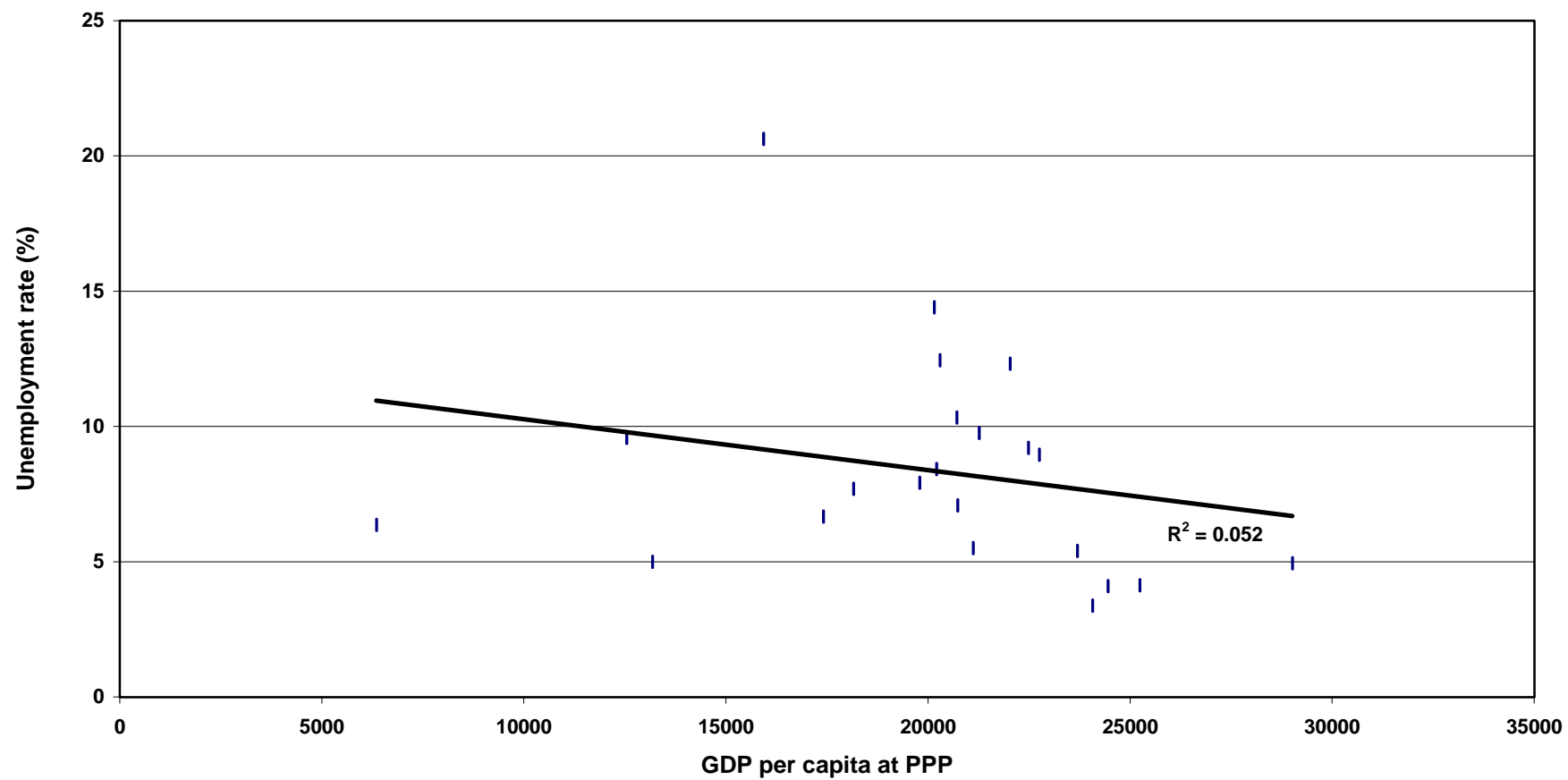
$R^2 = 0.6133$

Figure 20. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Transition Economies, latest year

$R^2 = 0.3638$
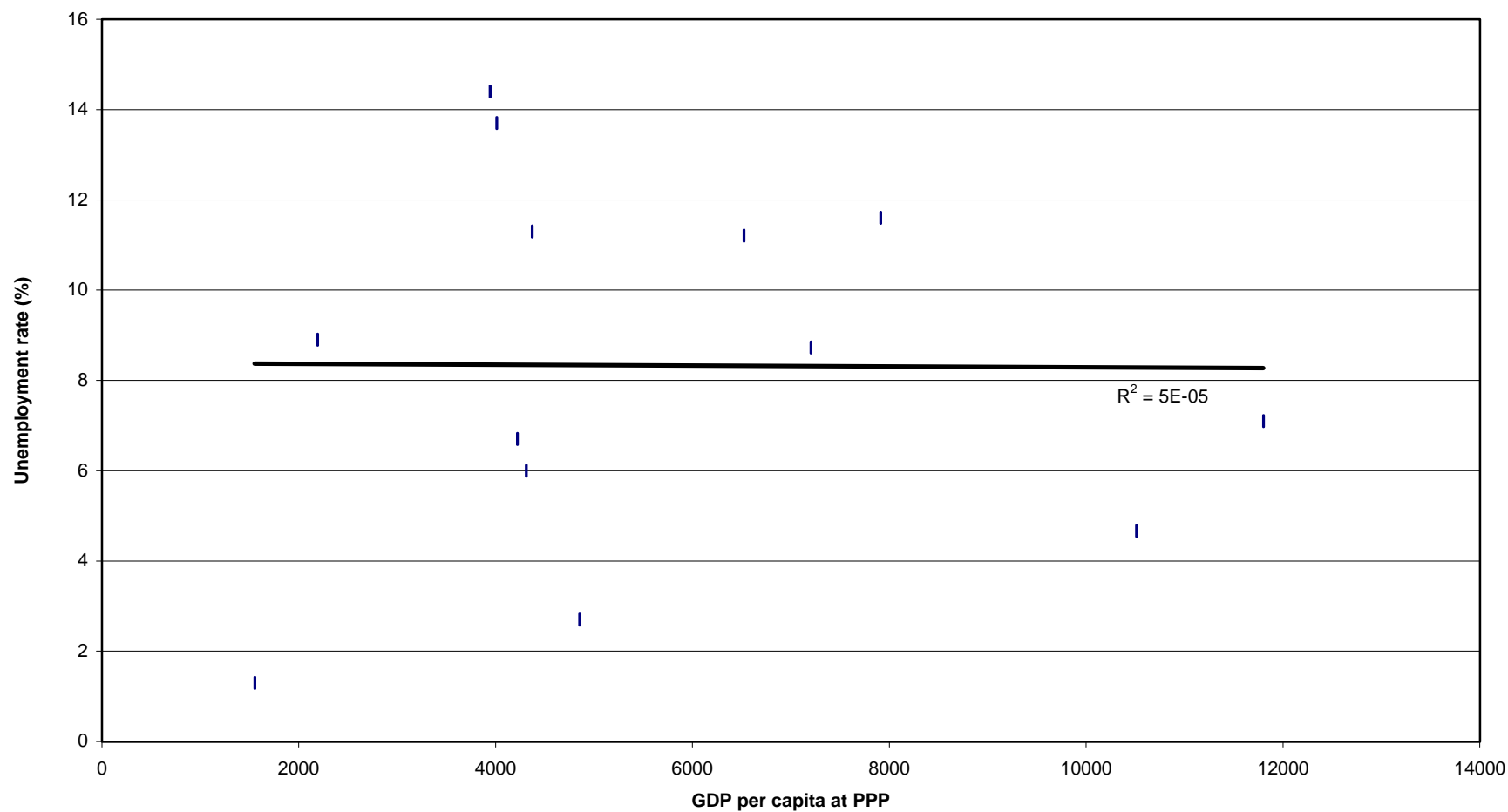
Y-axis: Employment in Services (%)

X-axis: GDP per capita at PPP

Figure 21. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Asia and the Pacific, latest year
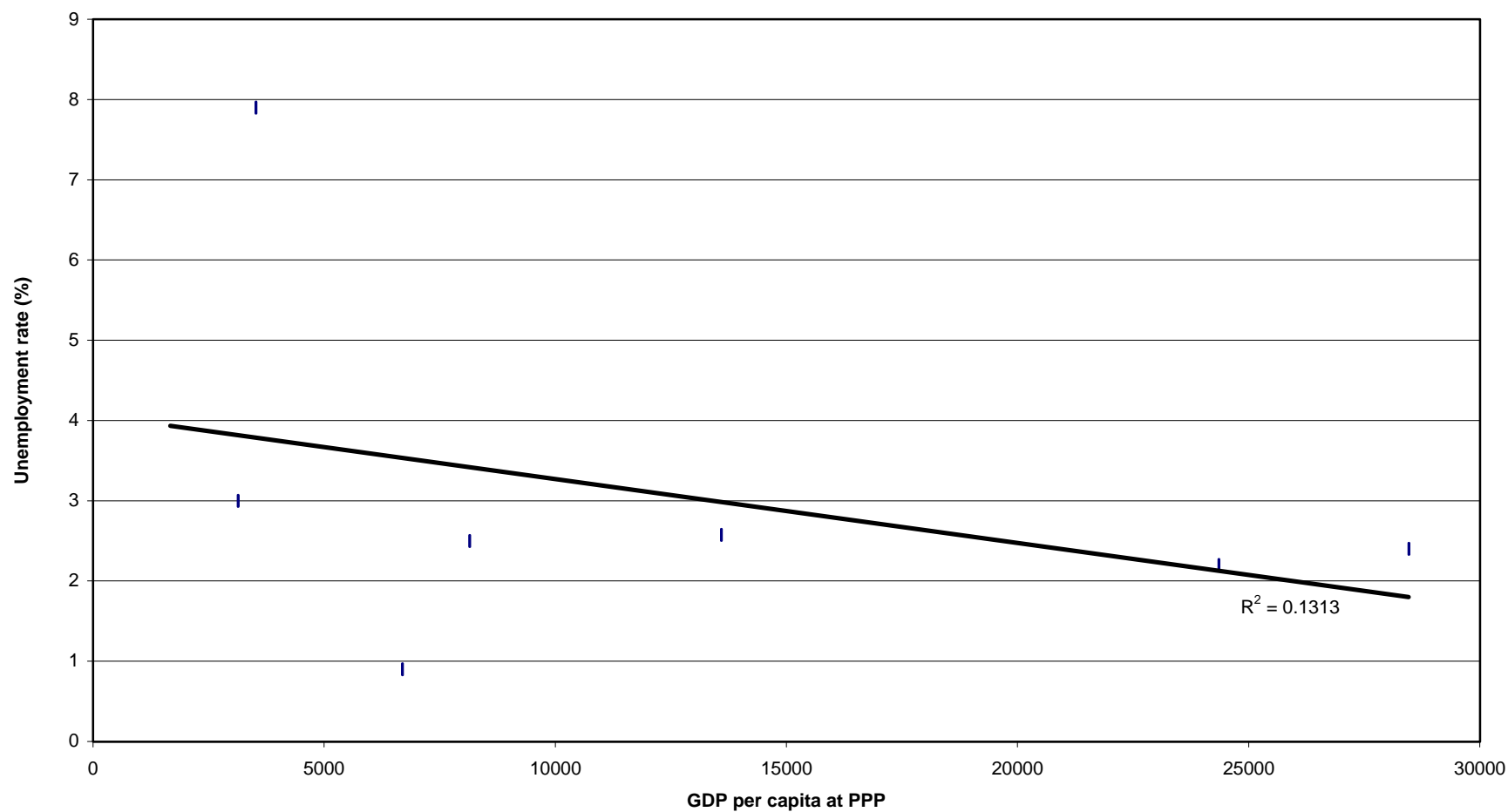
$R^2 = 0.7744$
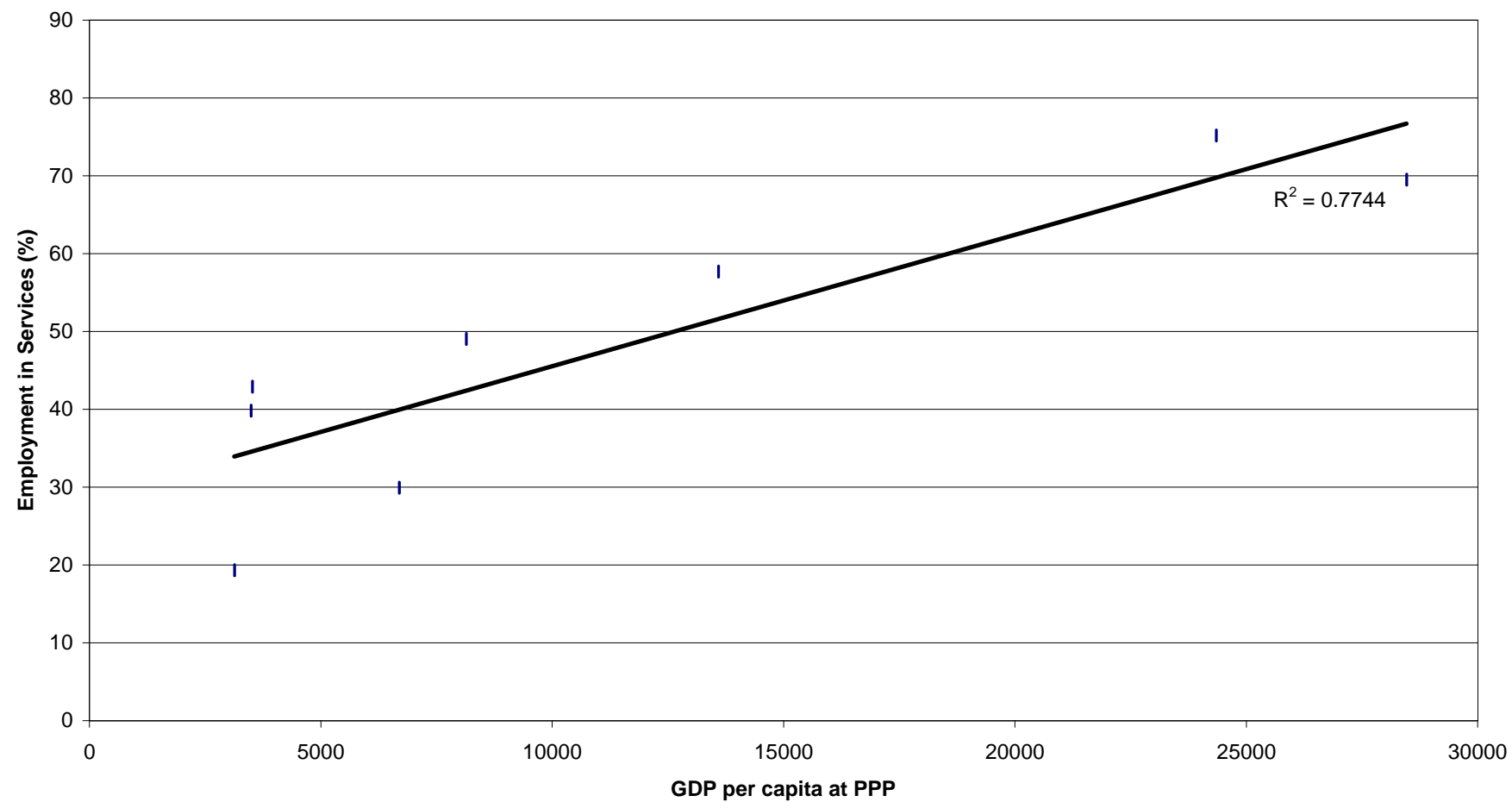
Figure 22. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Latin America and the Caribbean, latest year

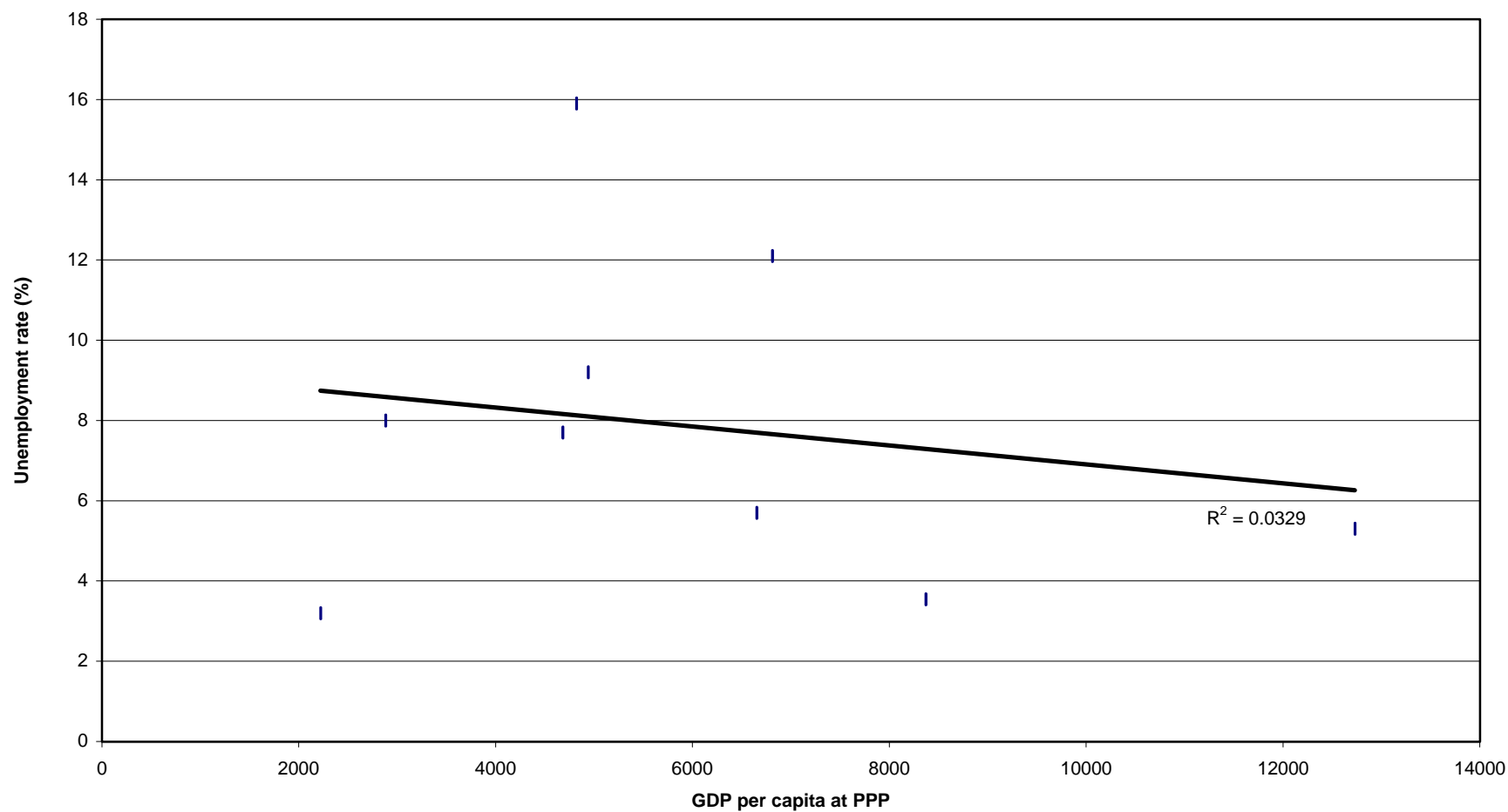$R^2 = 0.0438$

Employment in Services (%)

GDP per capita at PPP

Figure 23. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Developed (Industrialized) countries, latest year
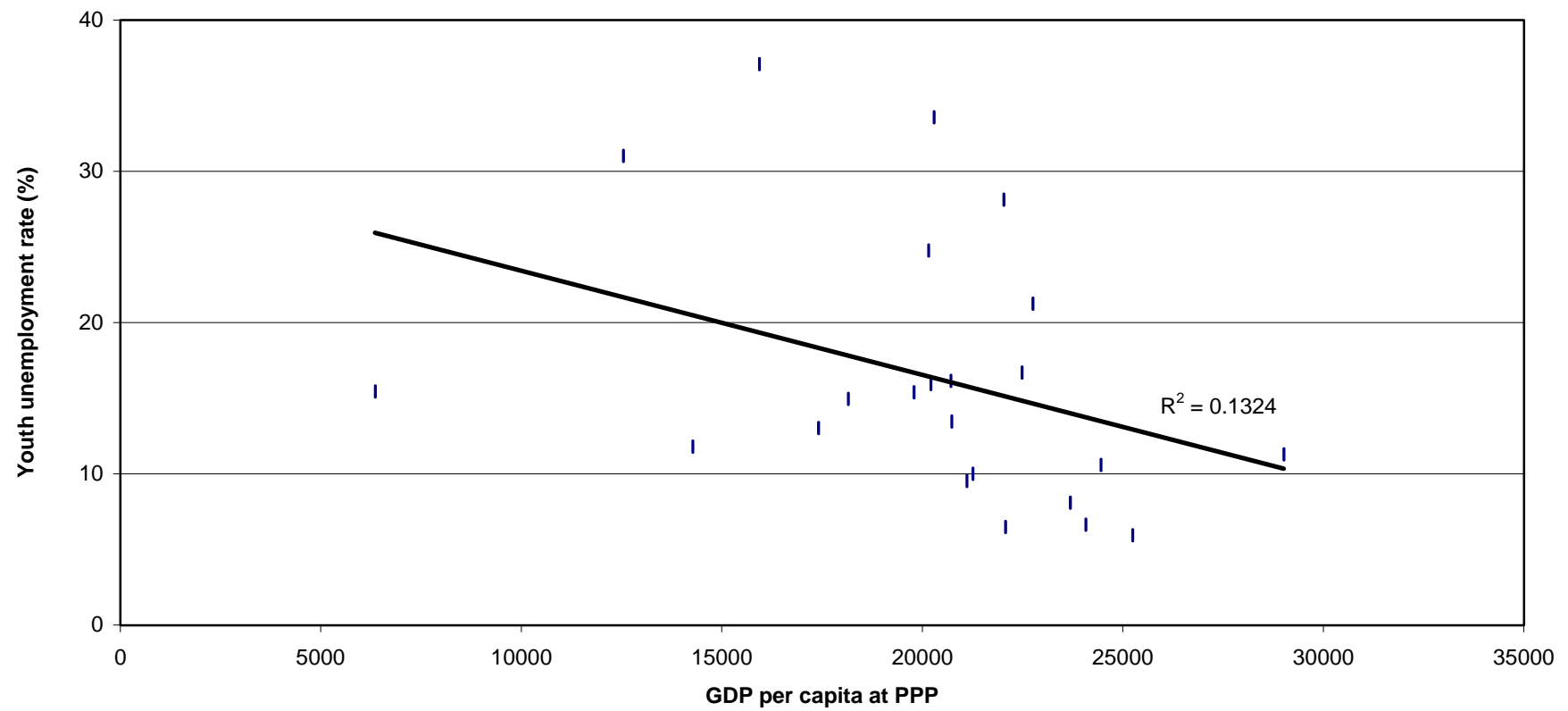
Figure 24. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Transition Economies, latest year

$R^2 = 0.5534$

Employment in Industry (%)

GDP per capita at PPP

Figure 25. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Asia and the Pacific, latest year

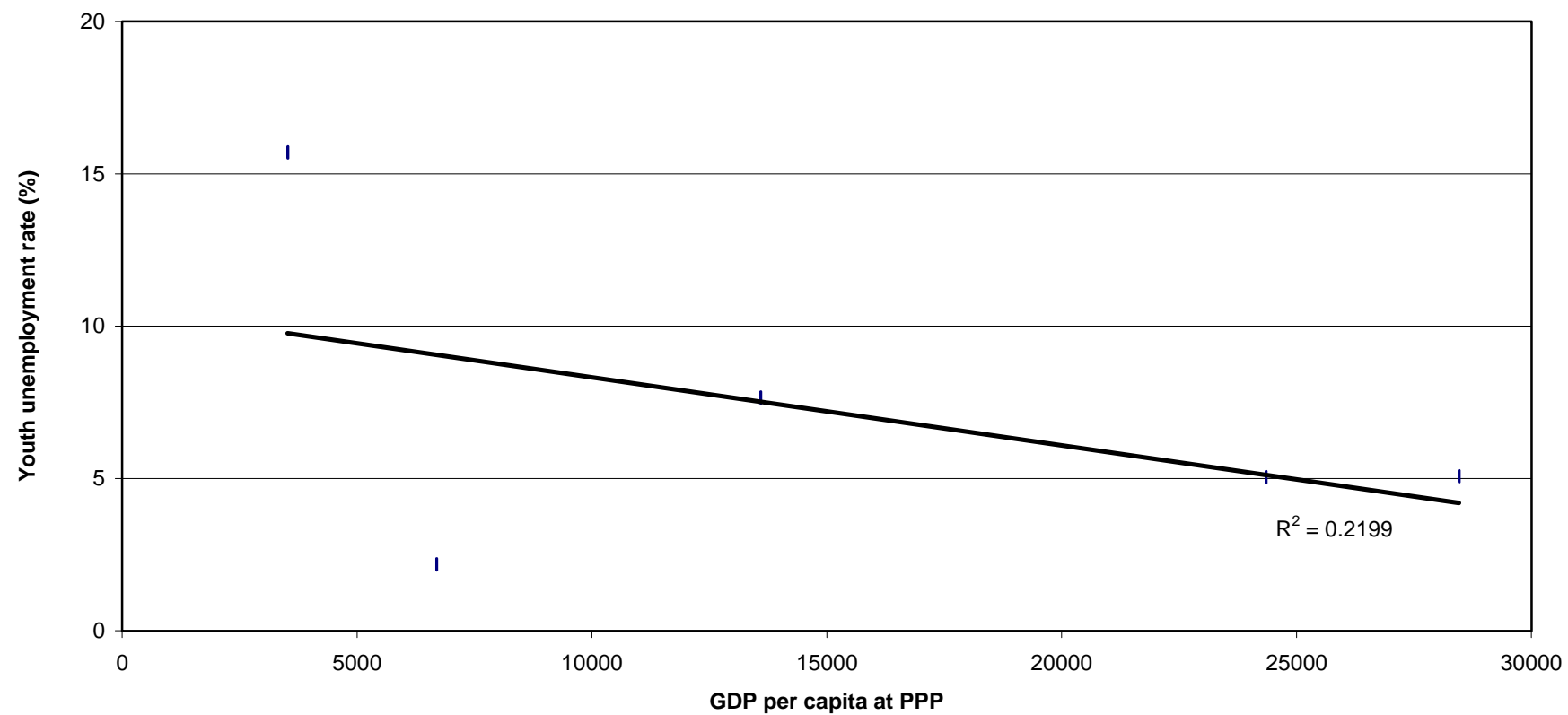$R^2 = 0.1904$

Employment in Industry (%)

GDP per capita at PPP

Figure 26. Employment by sector of both sexes and GDP per capita at puchasing power parity (PPP), Latin America and the Caribbean, latest year

$R^2 = 0.3896$

Employment in Industry (%)

GDP per capita at PPP

Figure 27. Unemployment rate of both sexes and GDP per capita at purchasing power parity (PPP), Developed (Industrialized) countries, latest year

$R^2 = 0.052$

Unemployment rate (%)

GDP per capita at PPP

Figure 28. Unemployment rate of both sexes and GDP per capita at purchasing power parity (PPP),
Transition Economies, latest year

Figure 29. Unemployment rate of both sexes and GDP per capita at purchasing power parity (PPP), Asia and the Pacific, latest year

$R^2 = 0.1313$

Figure 30. Unemployment rate of both sexes and GDP per capita at purchasing power parity (PPP),
Latin America and the Caribbean, latest year

Figure 31. Youth unemployment rate of both sexes and GDP per capita at purchasing power parity (PPP), Developed (Industrialized) countries, latest year

Youth unemployment rate (%)

GDP per capita at PPP

$R^2 = 0.1324$

Figure 32. Youth unemployment rate of both sexes and GDP per capita at purchasing power parity (PPP), Transition Economies, latest year

$R^2 = 0.4244$

Youth unemployment rate (%)

GDP per capita at PPP
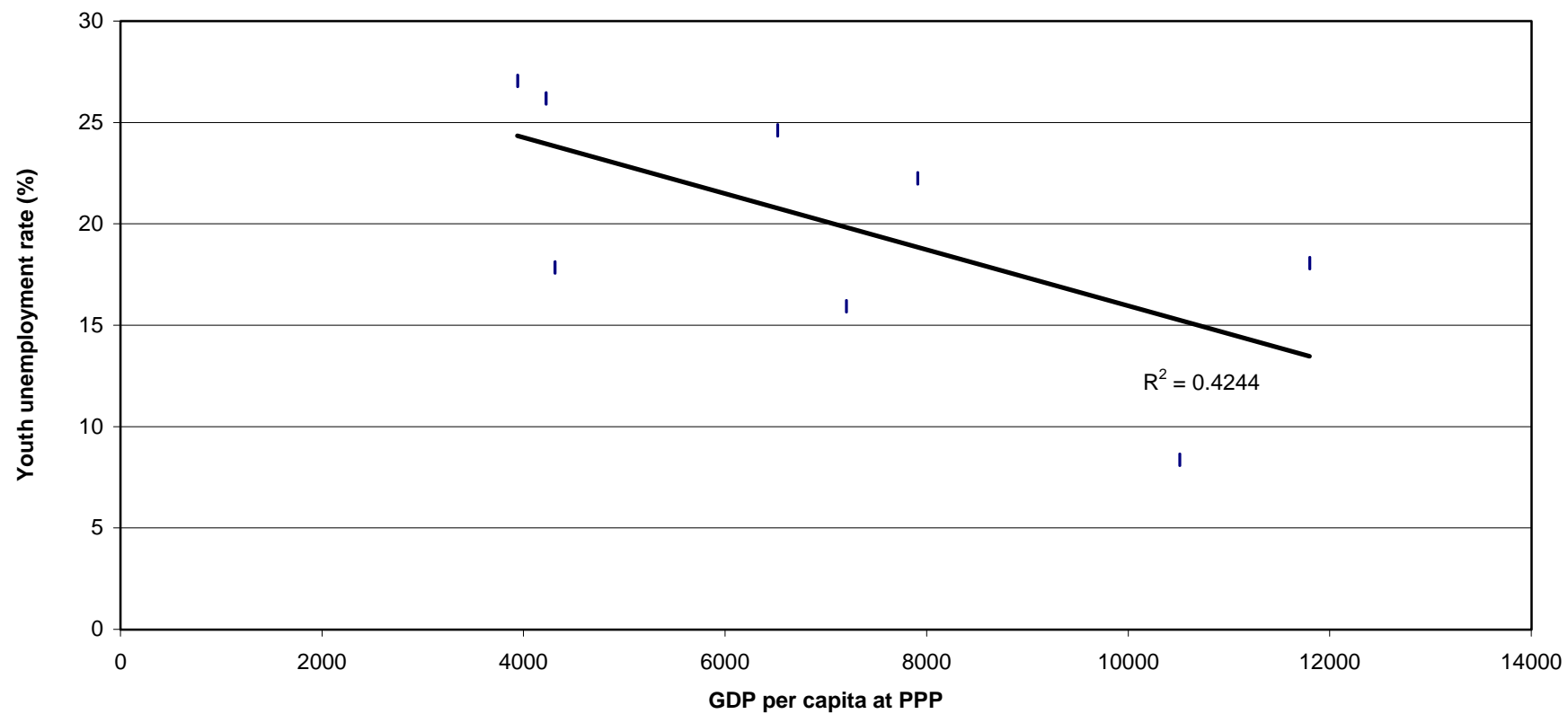
**Figure 33. Youth unemployment rate of both sexes and GDP per capita at purchasing power parity (PPP),
Asia and the Pacific, latest year**

$R^2 = 0.2199$

Youth unemployment rate (%)

GDP per capita at PPP

Figure 34. Youth unemployment rate of both sexes and GDP per capita at purchasing power parity (PPP), Latin America and the Caribbean, latest year

$R^2 = 0.0605$

Youth unemployment rate (%)

GDP per capita at PPP
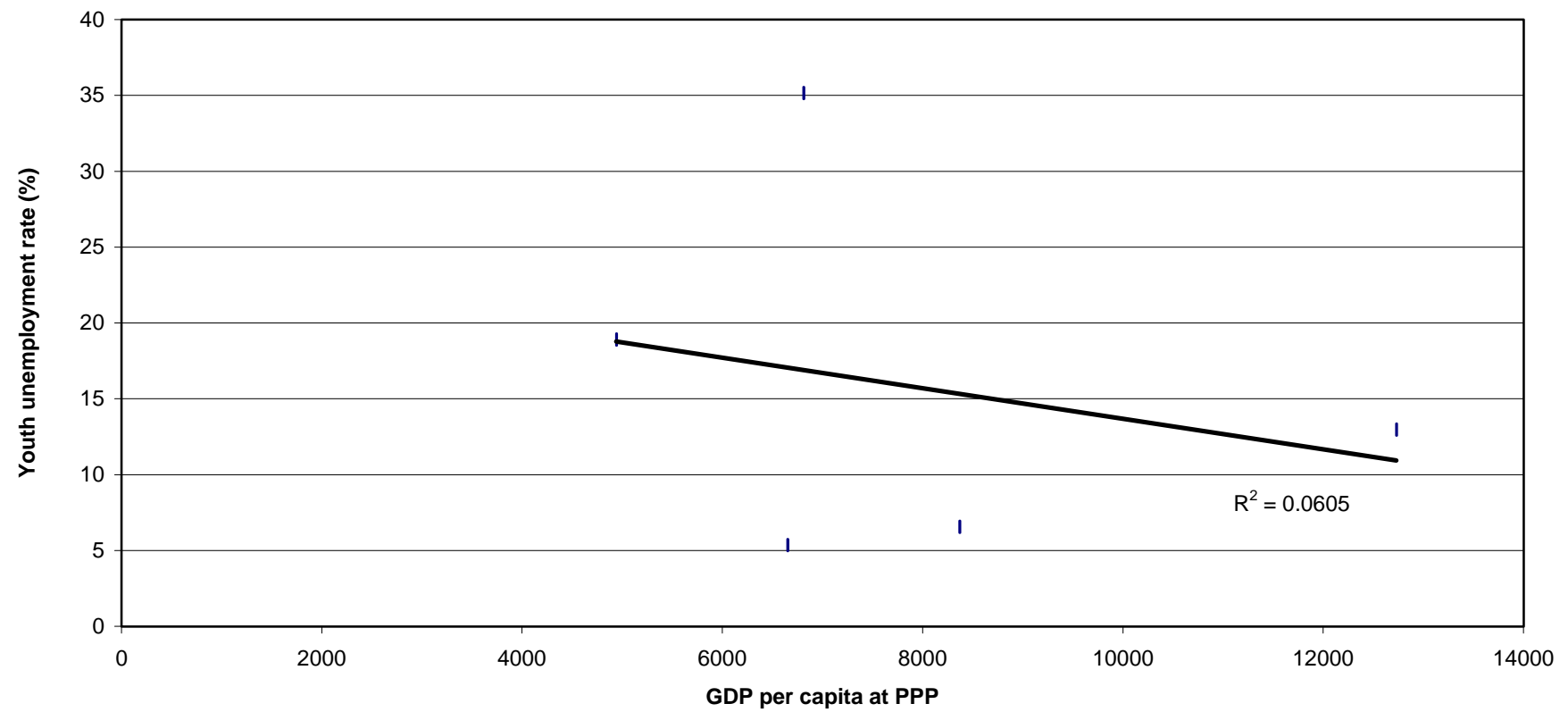
Figure 35. Labour force participation rate of male aged 15 years and over, and GDP per capita at purchasing power parity (PPP), latest year

$R^2 = 0.0519$
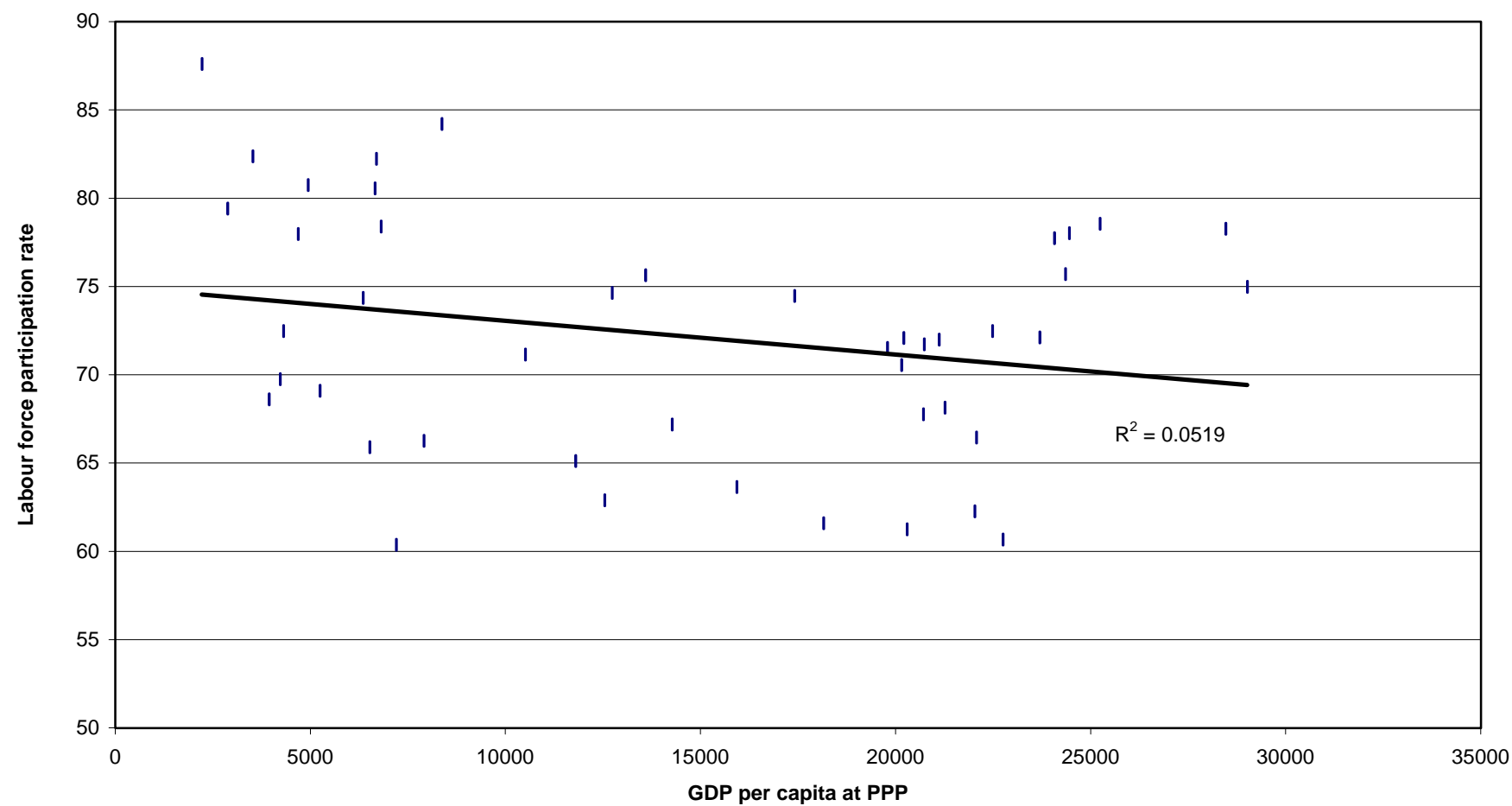
Figure 36. Labour force participation rate of female aged 15 years and over, and GDP per capita at purchasing power parity (PPP), latest year

$R^2 = 0.0047$

GDP per capita at PPP

Labour force participation rate
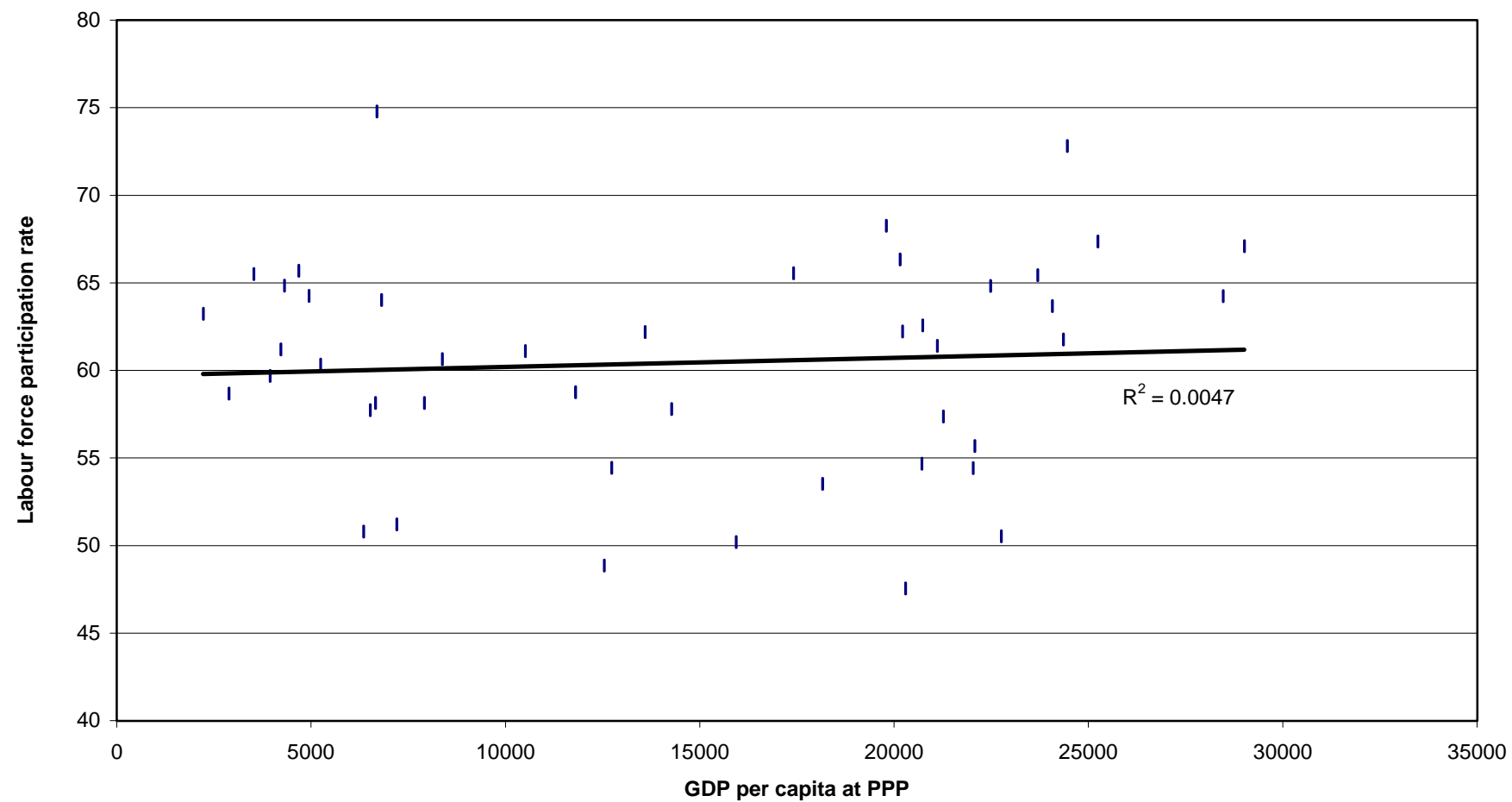
Figure 37. Employment-to-population ratio of males and GDP per capita at purchasing power parity (PPP), latest year

$R^2 = 0.0007$

Employment-to-population ratio

GDP per capita at PPP
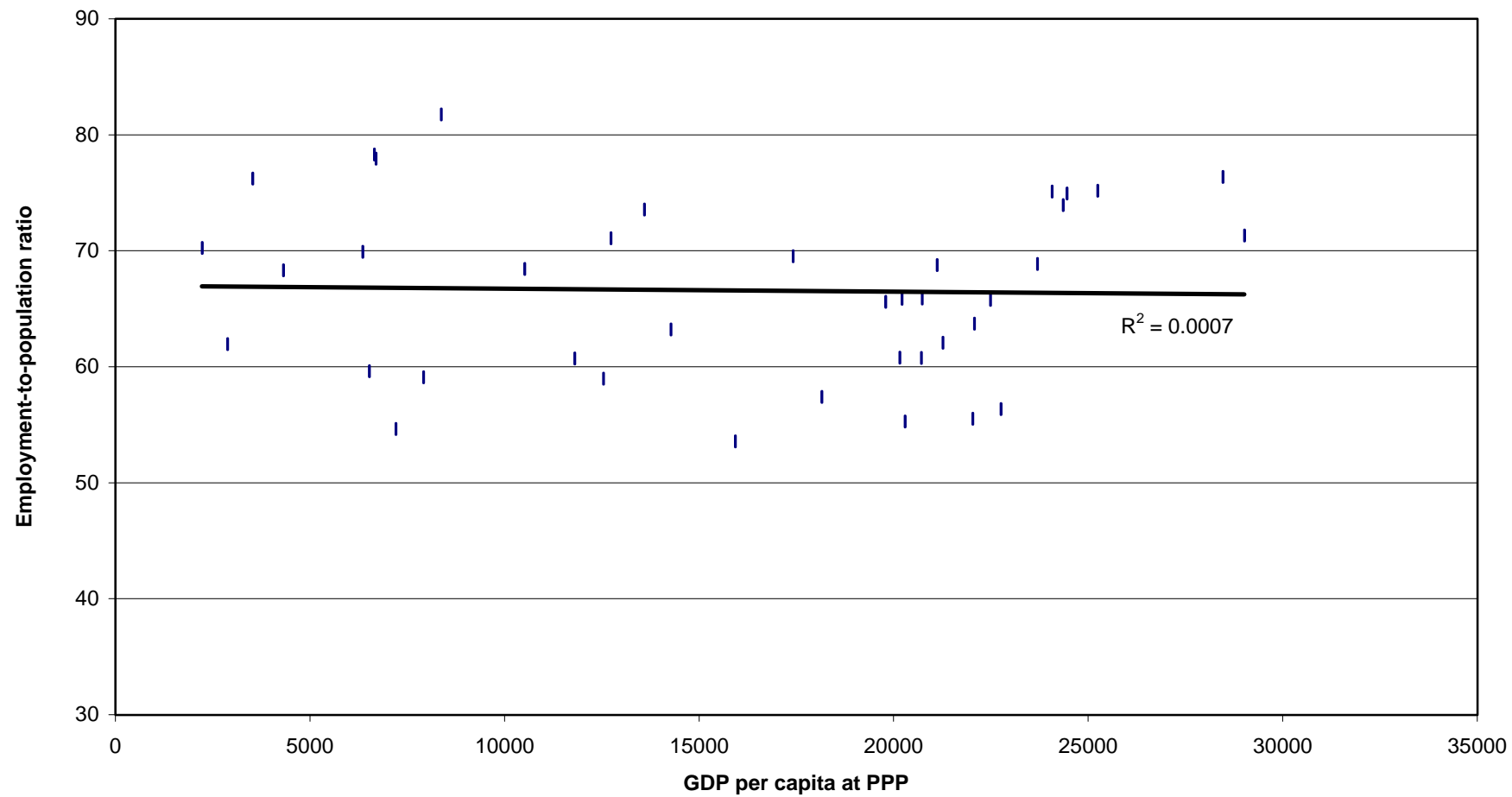
Figure 38. Employment-to-population ratio of females and GDP per capita at purchasing power parity (PPP), latest year

$R^2 = 0.1328$

Employment-to-population ratio

GDP per capita at PPP
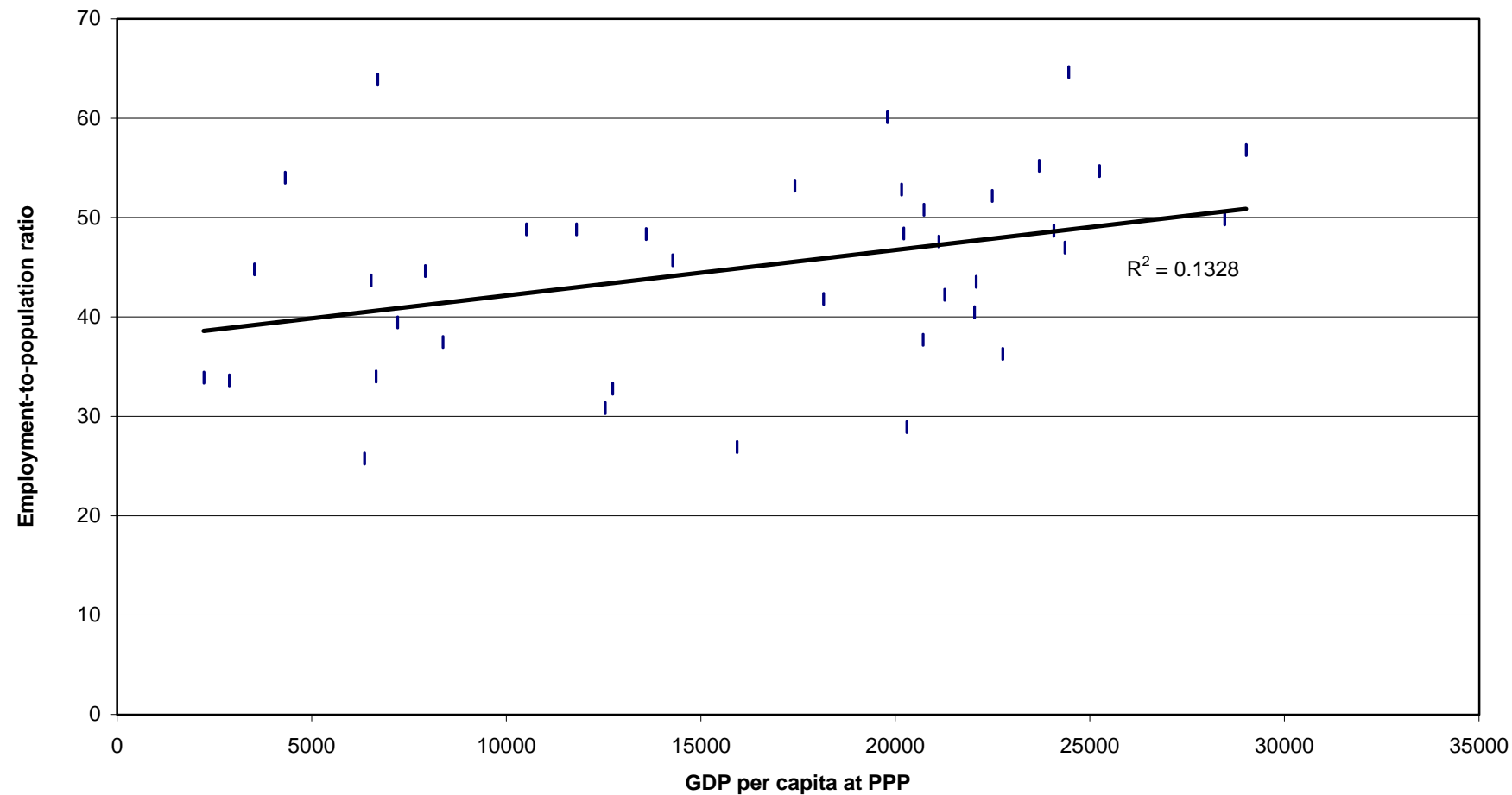
Figure 39. Employment by sector of males and GDP per capita at puchasing power parity (PPP), latest year
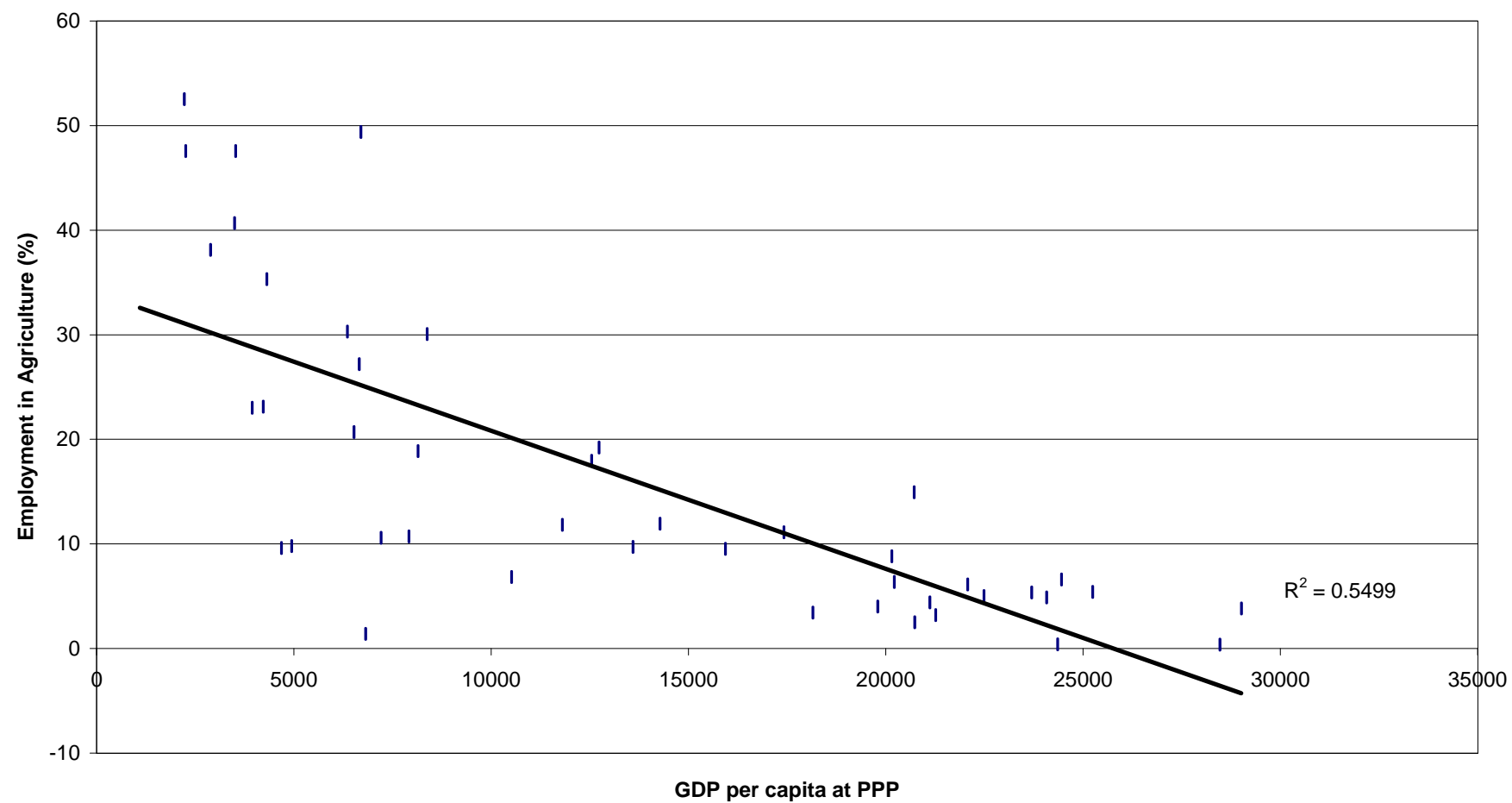
Figure 40. Employment by sector of females and GDP per capita at puchasing power parity (PPP), latest year
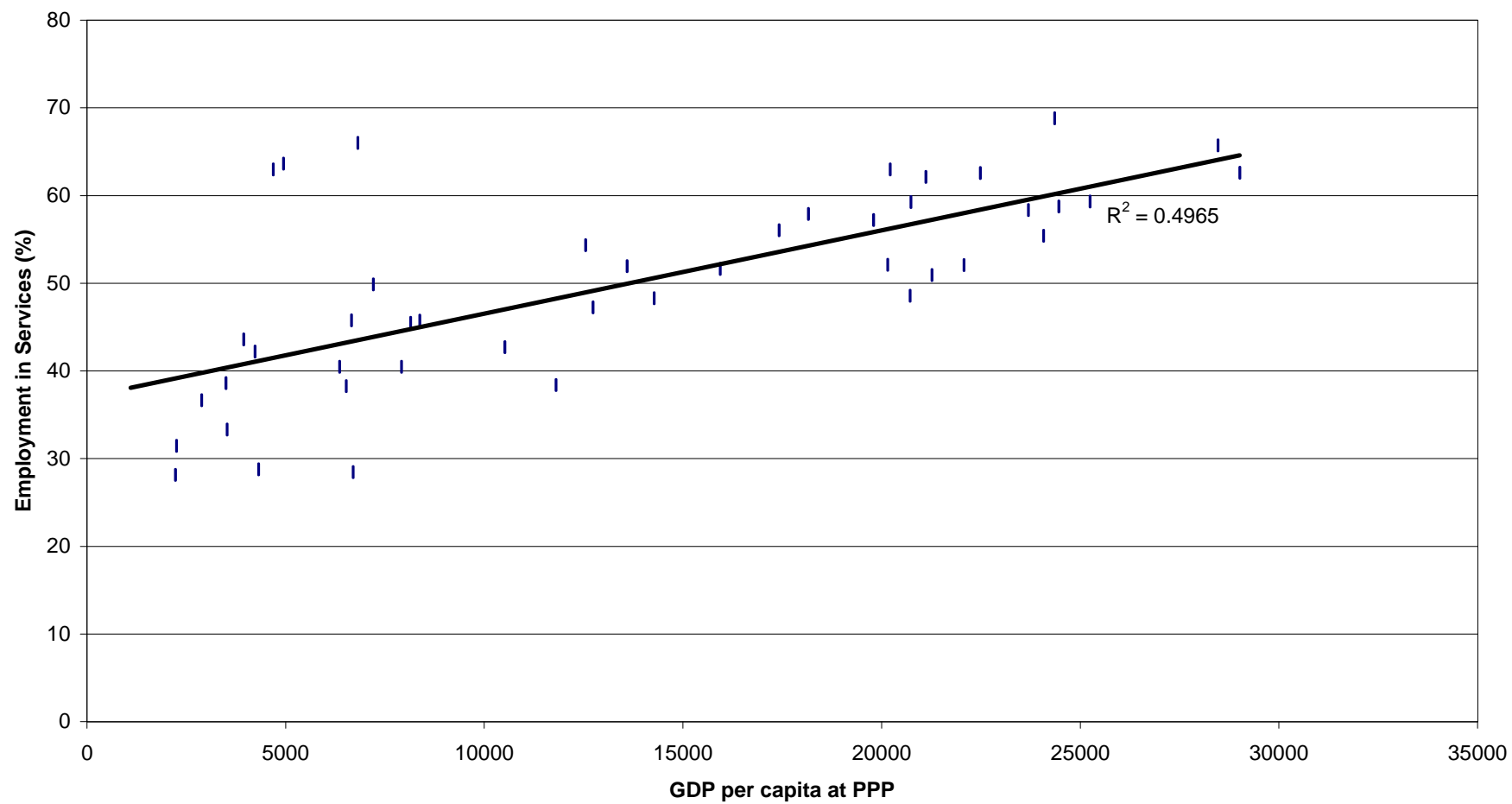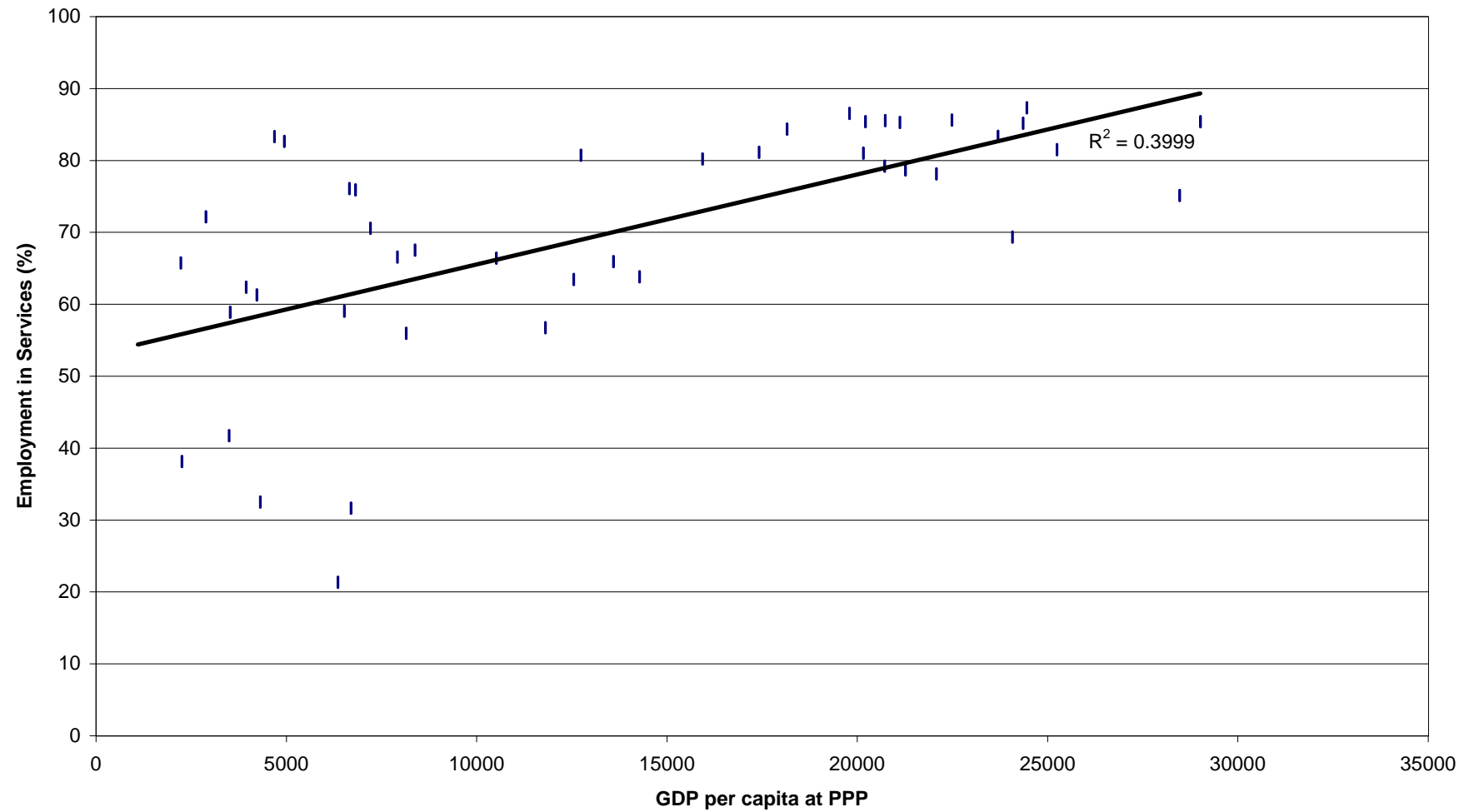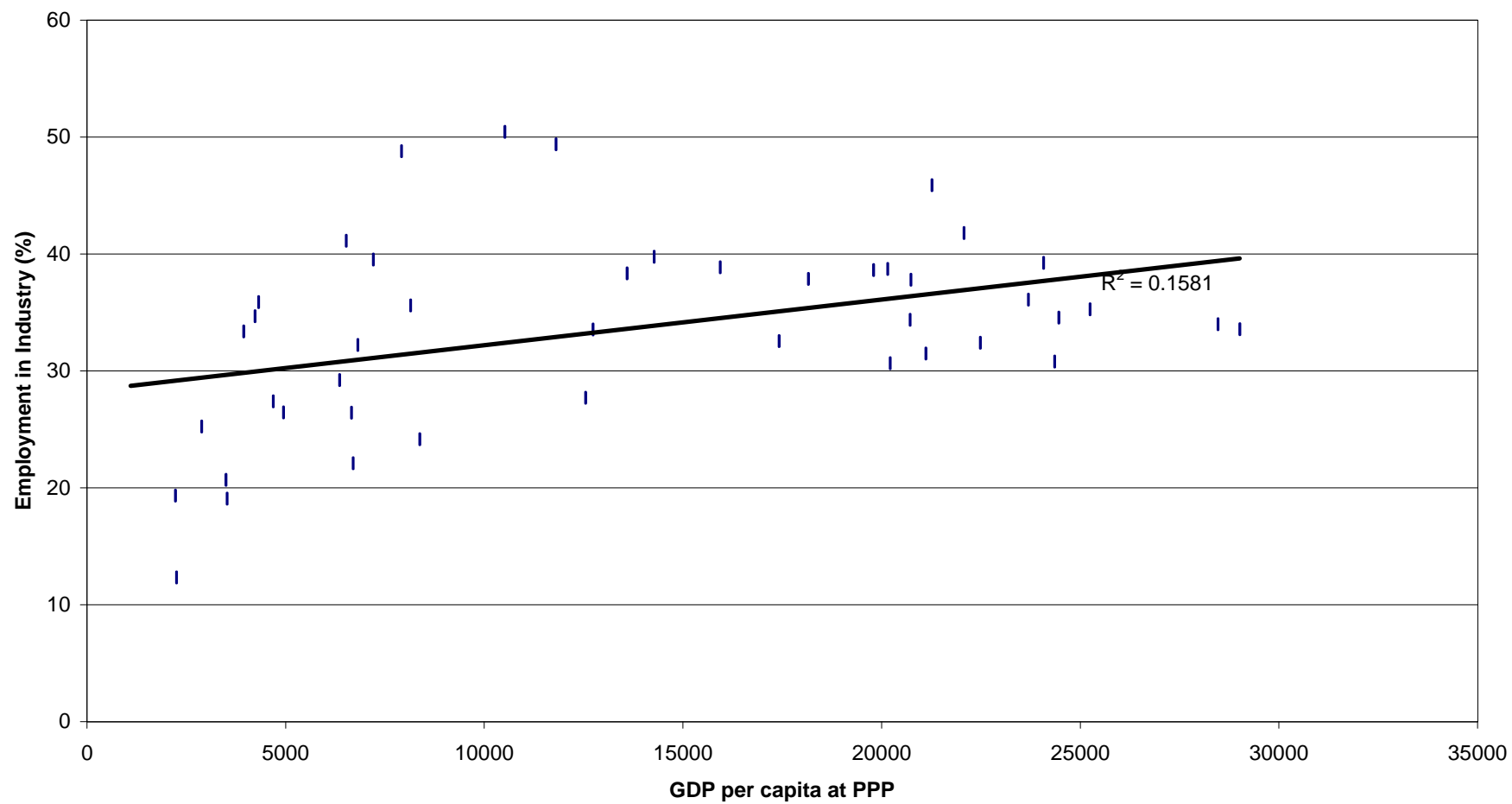
GDP per capita at PPP

Employment in Agriculture (%)

$R^2 = 0.2988$
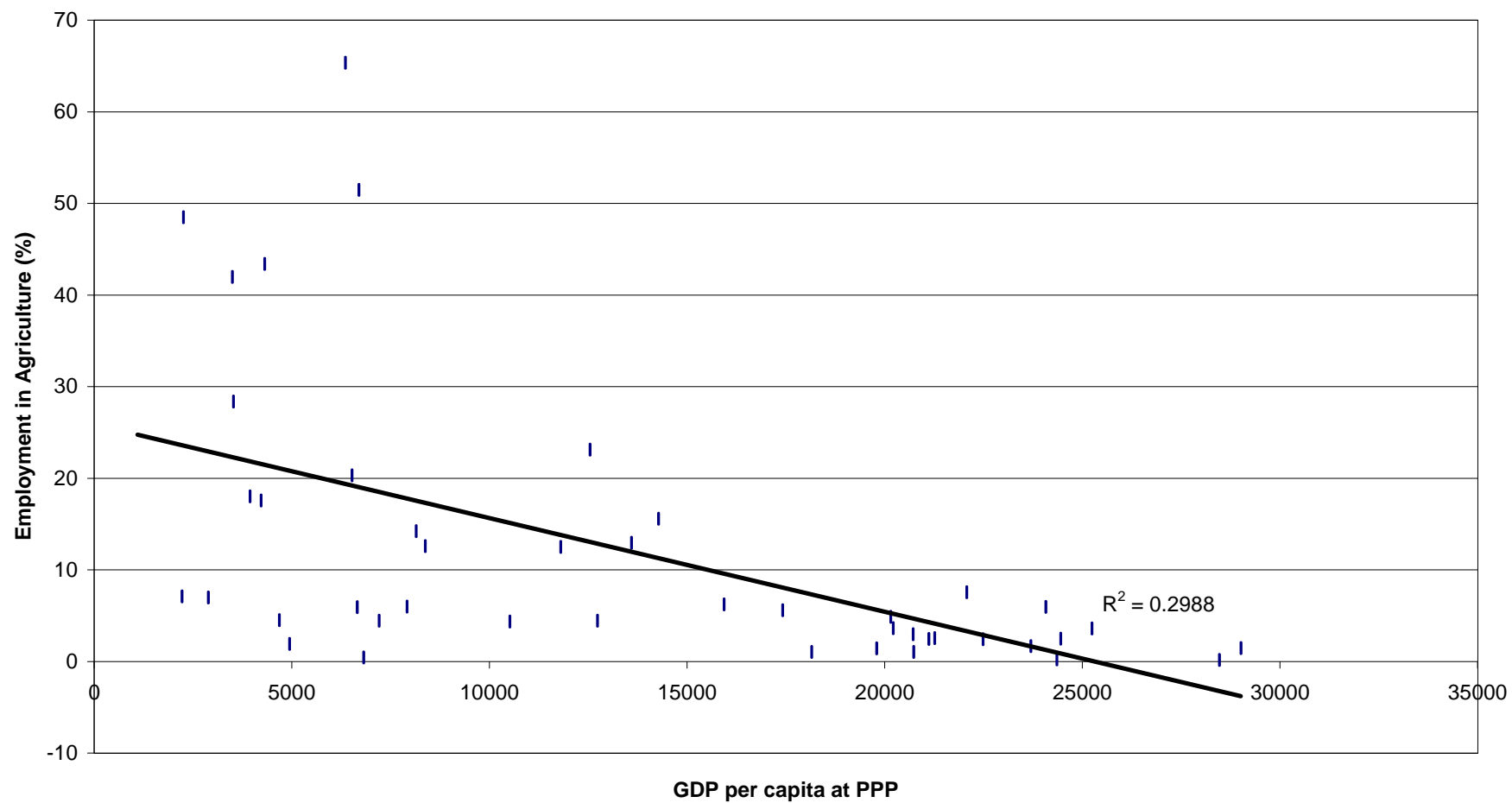
Figure 41. Employment by sector of males and GDP per capita at puchasing power parity (PPP), latest year

$R^2 = 0.4965$

Employment in Services (%)

GDP per capita at PPP

Figure 42. Employment by sector of females and GDP per capita at puchasing power parity (PPP), latest year

Figure 43. Employment by sector of males and GDP per capita at puchasing power parity (PPP), latest year

Figure 44. Employment by sector of females and GDP per capita at puchasing power parity (PPP), latest year
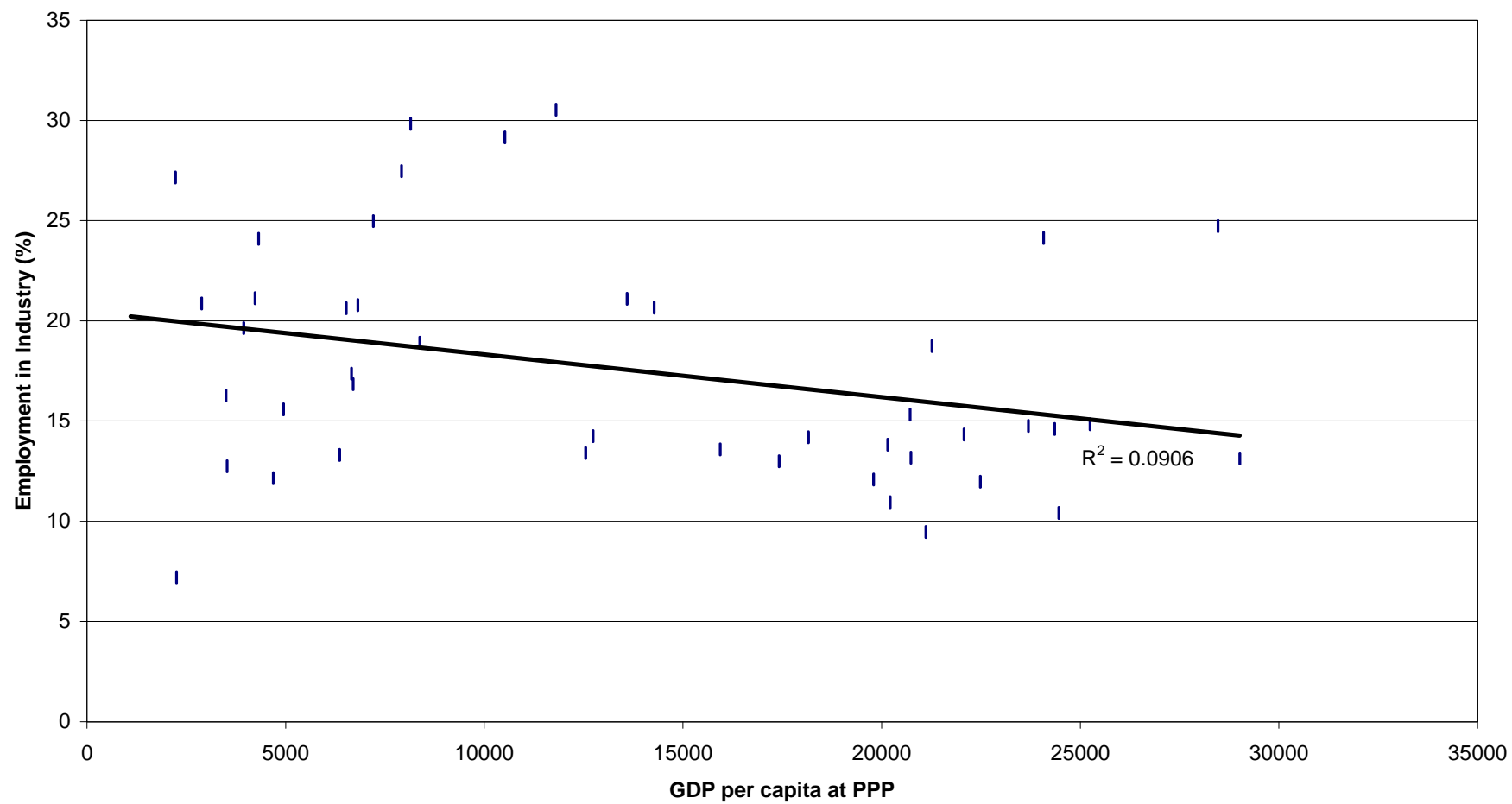
$R^2 = 0.0906$

Employment in Industry (%)

GDP per capita at PPP

Figure 45. Unemployment rate of male and GDP per capita at purchasing power parity (PPP), latest year

$R^2 = 0.0623$

Y-axis: Unemployment rate (%)
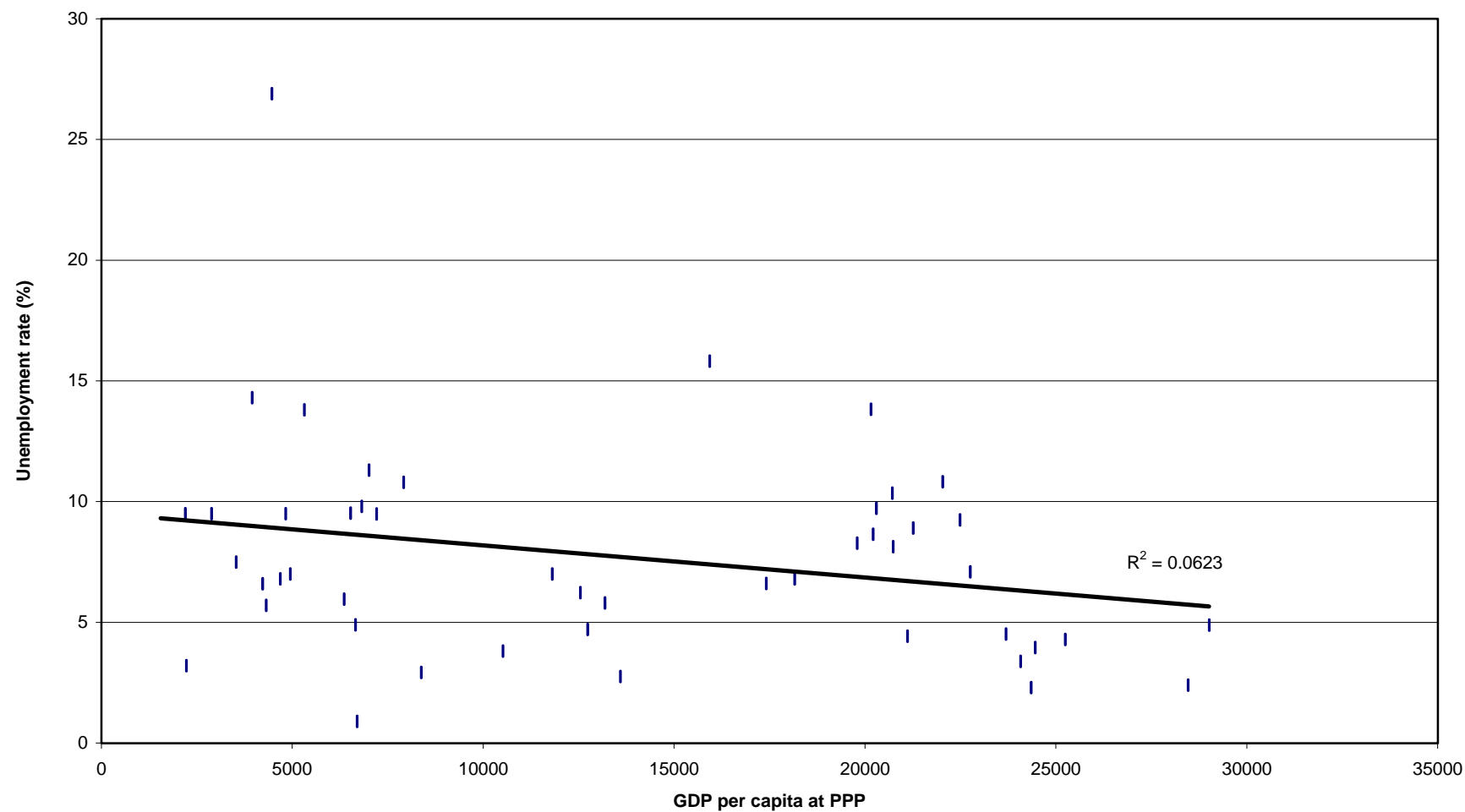
X-axis: GDP per capita at PPP

Figure 46. Unemployment rate of female and GDP per capita at purchasing power parity (PPP), latest year

Figure 47. Male youth unemployment rate and GDP per capita at purchasing power parity (PPP), latest year
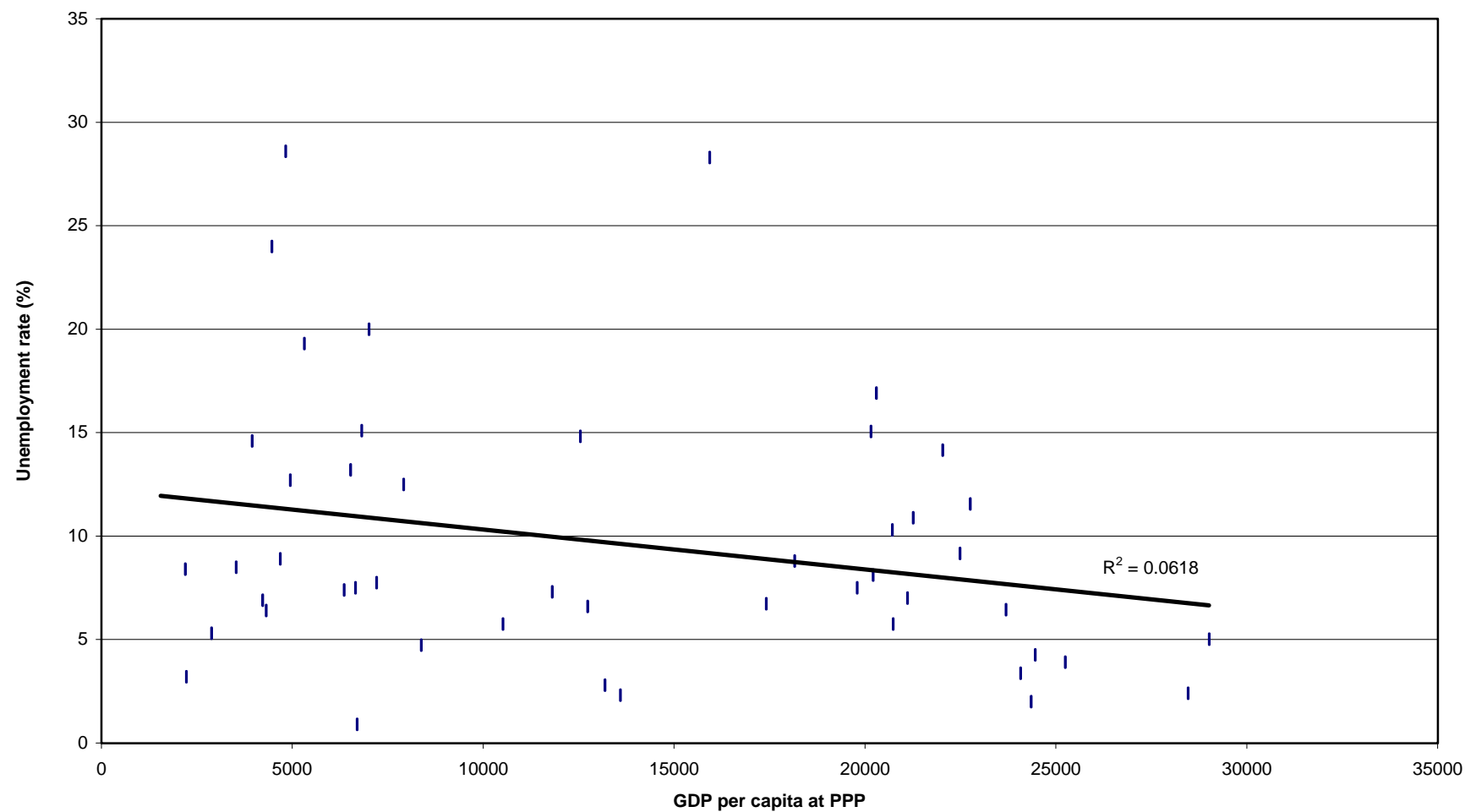
$R^2 = 0.0685$

Youth unemployment rate (%)
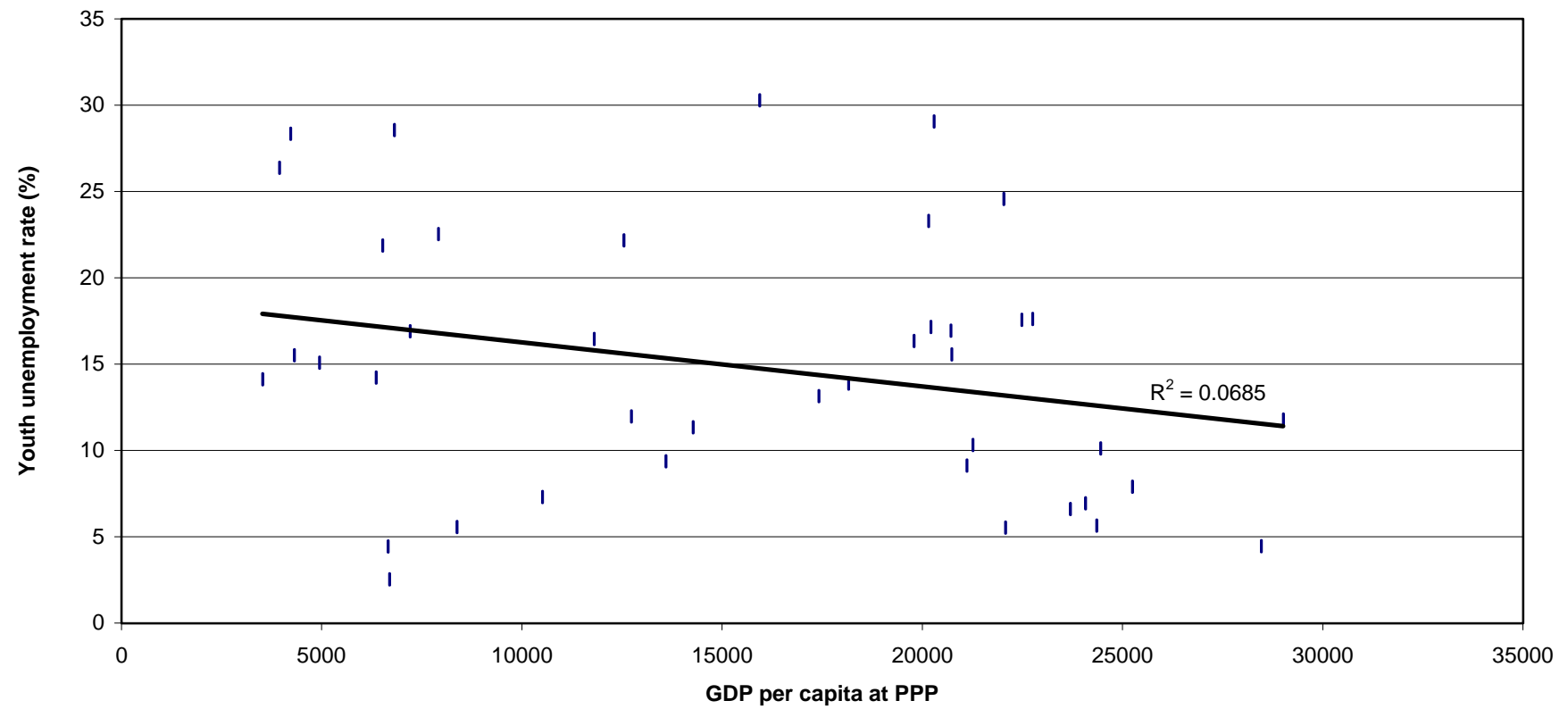
GDP per capita at PPP

Figure 48. Female youth unemployment rate and GDP per capita at purchasing power parity (PPP), latest year