The dynamics of socioeconomic inequality in health care utilisation in Europe: evidence from 8 waves of the ECHP

Teresa Bago d'Uva, Andrew M. Jones & Eddy van Doorslaer

Preliminary and incomplete

Abstract

This paper exploits European panel data to explore the evolution on inequalities and inequity in the use of health care over recent years. The data used in the analysis is taken from the *European Community Household Panel User Database* (ECHP-UDB). The ECHP was designed and coordinated by the Eurostat, and it was carried out annually between 1994 and 2001 (8 waves). We analyse health care utilisation over the previous year, represented by the number of visits to a GP and the number of visits to a specialist. The paper presents concentration curves, concentration indices (short-run and long-run), and mobility indices for the number of visits to a GP and to a specialist, by country across time. Additionally, we present concentration curves for the average number of visits across waves, using as ranking variable the average income across periods. Firstly, we analyse the inequalities for each country and wave. Secondly, we look into more detail at the long-run income-related inequalities, providing a comparison across countries. To compute indices of horizontal equity that exploit the panel data dimension of the ECHP we estimate latent class hurdle models, a model developed by Bago d'Uva (2005). A latent class (or finite mixture) framework is adopted in which individual effects are approximated using a discrete distribution. This framework offers an alternative representation of heterogeneity, where individuals are drawn from a finite number of latent classes.

Keywords: Inequality, inequity, health care utilisation, panel data, ECHP

Acknowledgements: The *European Community Household Panel Users' Database*, version of December 2003, was supplied by Eurostat. This paper derives from the project "The dynamics of income, health and inequality over the lifecycle" (known as ECuity III Project), which is funded in part by the European Commission's Quality of Life and Management of Living Resources programme (contract QLK6-CT-2002-02297).

1. Introduction

Recent health economics literature has devoted a great deal of attention to the measurement of socioeconomic inequality and inequity the utilisation of health care in European countries. In particular, Van Doorslaer et al (2002) and van Doorslaer, Koolman and Jones (2004) have provided cross-country comparisons using data from the 1996 wave of *European Community Household Panel User Database* (ECHP-UDB).

This paper presents new evidence on inequalities and inequities in health care utilisation in Europe. We exploit all 8 waves of the ECHP, corresponding to the period 1994 to 2001. van Doorslaer et al (2002) and van Doorslaer, Koolman and Jones (2004) have not included Finland in their comparative analyses due to unavailability of data for this country. ECHP data for Finland is now available, allowing us to include this country in our analysis. The major contributions of this study arise from the fact that we are now able to exploit the full ECHP dataset, in particular, its panel structure. Regarding the analysis of income-related inequality in health care use, we complement the standard measures for each cross-section with more reliable long-run measures, making use of the information on income and health care throughout the observed periods.

In order to calculate indices of horizontal inequity that control for unequal need distributions, we estimate latent class panel data models. The panel feature of the data makes it possible to control for individual unobserved heterogeneity. As noted by Riphahn et al (2003), it is important to account for this type of heterogeneity because some unobserved individual specific characteristics may generate significant influences on health care demand. Attitudes towards health care, preferences, risk aversion, as well as genetic frailty and morbidity are some of the unobserved factors that influence health care use. Despite the importance of accounting for individual unobserved heterogeneity, this is seldom done in empirical modelling of health care utilisation. With one notable exception (Van Ourti, 2004), the literature on health care inequity has always used cross-sectional methods. Van Ourti (2004) has developed a random effects hurdle model which he has used to produce horizontal inequity indices for Belgium. Despite the advantages of panel data methods to model health care utilisation, their use for the computation of inequity indices poses further challenges. In particular, the prediction of need-expected levels of utilisation calculation is not straightforward. We present here a discussion of different possible procedures to calculate those predictions. This study will ultimately assess the extent to which panel data methods can lead to different conclusions regarding inequalities and inequities in health care use across European countries.

2. Data

The data used in the analysis presented here is taken from the *European Community Household Panel User Database* (ECHP-UDB). The ECHP was designed and coordinated by the Eurostat, and it was carried out annually between 1994 and 2001 (8 waves). This survey contains socioeconomic, demographic, health and health care utilisation variables, for a panel of individuals aged 16 or older. The data result from a standardised questionnaire, which allows for cross-country comparisons as well as longitudinal analysis. We use all the information that is available for 13 EU member states: Austria, Belgium, Denmark, Finland, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain and the United Kingdom. In the United Kingdom, Luxembourg and Germany, the ECHP was carried out from 1994 to 1997 (waves 1 to 3), after which it was replaced by national panel surveys. Austria joined the survey in 1995 (wave 2) and in Finland it started only in 1996 (wave 3).

We analyse health care utilisation over the previous year, represented by the number of visits to a GP and the number of visits to a specialist. These data are available from wave 2 onwards (in wave 1, the information is not detailed by type of doctor). The ECHP income variable is total net household income. We use this variable deflated by PPPs and national CPIs, in order to allow for comparability cross-country and across waves. The income variable was further divided by the OECD modified equivalence scale in order to account for household size and composition. The sample used is an unbalanced panel of individuals observed for up to 6 waves in the case of Finland, 1 or 2 waves in the United Kingdom, Germany and Luxembourg and up to 7 waves for the remaining countries. The descriptive analysis presented here uses cross-sectional individual weights.

Tables 1 to 3 contain the averages by country and wave of the three variables used in the measurement of socioeconomic inequality in health care: equivalised household income, specialist visits and GP visits. The highest levels of income are observed for Luxembourg. The countries with the lowest income levels are Portugal and Greece, followed by Spain and Italy. In general, there was an increase in the income levels throughout the panel, especially for Ireland (31%), Spain (28%) and Portugal (24%). Table 2 shows that the levels of utilisation of specialist care vary considerably across Europe. Ireland is the country with the lowest average utilisation throughout the observed years, followed by Finland and Denmark which have a similar pattern. Germany has the highest number of specialist visits, on average. There is also large variation in the average number of GP visits (see Table 3) observed across countries, with the lowest values for Finland and Greece, while Belgium and Germany have the highest values (as well as Italy, towards the end of the period, and Austria, especially in the beginning and the end of the period).

In the analysis of inequity in health care, we use additional variables to represent need and and other non-need factors. We use one-year lagged health measures based on two questions: (a) responses to a question on self-assessed general health status as either very good, good, fair, bad or very bad; and (b) responses to "Do you have any chronic physical or mental health problem, illness or disability? (yes/no)" and, if so, "Are you hampered in your daily activities by this physical or mental health problem, illness or disability? (no; yes, to some extent; yes, severely)". We use two dummies to indicate either some limitation or severe limitation. Gender and age are represented by dummy variables: m30-44, m45-59, m60-69, m70+, f16-29, f30-44, f45-59, f60-69 and f70+. Apart from income, the following non-need variables are considered: (i) the highest level of general or higher education completed, i.e. recognised third level education (ISCED 5-7), second stage of secondary level of education (ISCED 3) or less than second stage of secondary education (ISCED 0-2)); (ii) Marital status, distinguishing between married, separated/divorced, widowed and unmarried (including cohabiting); (iii) number of children in the household (categorised as:

under 5, aged 5 to 11 and aged 12 to 18); (iv) Activity status includes employed, self-employed, student, unemployed, retired, doing housework and 'other economically inactive'. We have also included indicators of region of residence (EU's NUTS 1 level, Nomenclature of Statistical Territorial Units) for the countries for which such information was made available.

3. Methods for measurement of inequality

3.1 Cross-section measures

This section presents the methods for the measurement of income-related inequality in the utilisation of health care (measured by GP visits and specialist visits) for the 13 EU member states. The analysis uses concentration curves (Wagstaff et al, 1991; Kakwani et al, 1997) and concentration indices (Wagstaff et al, 1991; Kakwani et al, 1997; Lambert, 1993). It should be noted that the concentration indices and the concentration curves represent income-related inequality in health care, thus not taking into account the need for such care. The analysis of inequity in health care utilisation, controlling for unequal distribution of need, is covered below.

In our application, the concentration curve is a plot of the cumulative percentage of the number of doctor visits against the cumulative percentage of the sample, ranked in increasing order of equivalised household income. If the same number of visits is observed for all the individuals in the sample, the concentration curve is simply a 45º line through the origin. A curve lying above the 45º line (equality line) represents a pro-poor distribution of health care utilisation (i.e. health care is more concentrated amongst the poor). Similarly, a pro-rich distribution of health care utilisation leads to a concentration curve lying below the equality line.

We use the concentration index as a summary measure of income-related inequality in doctor visits. The concentration index is defined as twice the area between the concentration curve and the line of equality. Thus, a value of zero for the concentration index means that there is no income-related inequality in the distribution of health care utilisation. The concentration index takes on positive (negative) values when there is pro-rich (pro-poor) inequality, which corresponds to the concentration curve lying below (above) the equality line. We compute the concentration index of the number of visits in wave *t*, CI^t, using the convenient covariance formula (for example, Kakwani, 1980):

$$CI^t = \frac{2}{\bar{y}_t}\operatorname{cov}\left(y_{it}, R_i^t\right) \quad (1)$$

where y_{it} is the number of visits to a doctor for individual *i* in period *t*, $\bar{y}_t$ is the average of y_{it} across individuals in period *t*, and R_i^t is the relative rank of individual *i* in the distribution of income in period *t*.

3.2 Longitudinal measures

The analysis of income-related inequality in health care utilisation is extended to the analysis of long-run inequality, exploring the panel feature of the data. This follows the methodology proposed by Jones and Lopéz-Nicolás (2004) for the analysis of income-related inequality in health. Similarly to what was done for health in Jones and Lopéz-Nicolás (2004), we define the long-run concentration index of health care utilisation, CI^T, as the concentration index for the average number of visits across periods, using as ranking variable the average income across periods. Jones and Lopéz-Nicolás (2004) show that, when the income ranking remains constant across time, the long-run concentration index equals the (weighted) average of the short-run concentration indices. These two differ to the extent that income ranks change over time and those changes are associated with systematic differences in health care utilisation. Jones and López-Nicolás (2004) provide a simple measure of how much long-run inequality differs from what can be obtained using repeated cross-sections. The index of health-related income mobility is defined as one minus the ratio between the long-run concentration index and the weighted average of the cross sectional concentration indices:

$$MI^T = 1 - \frac{CI^T}{\sum_t w_t CI^t}, \quad \text{where } w_t = \frac{\bar{y}^t}{T\bar{\bar{y}}^T}, \tag{2}$$

and $\bar{\bar{y}}^T$ equal to the average of $\bar{y}^t$ across t.

Section 5 below presents concentration curves, concentration indices (short-run and long-run), and mobility indices for the number of visits to a GP and to a specialist, by country across time. Additionally, we present concentration curves for the average number of visits across waves, using as ranking variable the average income across periods. These will be referred to as long-run concentration curves. Firstly, we analyse the inequalities for each country and wave. Secondly, we look into more detail at the long-run income-related inequalities, providing a comparison across countries.

3.3 Measurement of horizontal inequity

The concentration index measures income-related inequality in health care. This is not the same thing as inequity in health care. For example, variations in the use of health care that are attributable to differences in morbidity may be seen as unavoidable and hence legitimate sources of inequality (see e.g., van Doorslaer, Koolman and Jones, 2004). Usually, the horizontal version of the egalitarian principle is interpreted to require that people in equal need of care are treated equally, irrespective of characteristics such as income, place of residence, race, etc. While the concentration index of medical care use (C_M) measures the degree of inequality in the use of medical care by income, it does not yet measure the degree of inequity. For any inequality to be interpretable as inequity, legitimate or need-determined inequality has to be taken into account.

There are two broad ways of standardising distributions for differences in need: the direct and the indirect method. The direct method proceeds by computing a concentration index for medical care use that would emerge if each individual had the same need characteristics as the population as a whole. Wagstaff *et al.* (1991) have used this procedure to compute what they call HI_{WVP} indices, which are essentially directly standardised concentration indices. More recently, Wagstaff and Van Doorslaer (2000) have advocated the technique of indirect standardisation for the measurement of so-called HI_{WV} indices on the grounds that it is computationally easier and does not rely on grouped data. A measure of the need for medical care is obtained for each individual as the predicted use from a regression on need indicators. This means that in order to statistically equalize need for the groups or individuals to be compared, one is effectively using the average relationship between need and treatment for the sample as a whole as the vertical equity norm and horizontal inequity is measured by systematic deviations from this norm by income level.

The issue of the role of explanatory models in the measurement of inequity deserves some further attention. Recently, some authors have drawn attention to the potential biases involved in these standardisation procedures. First, the problem of determining which systematic variations in medical care use by income are "needed" and therefore, in a sense, justifiable, and which are not, bears some resemblance to the problem of determining legitimate compensation in the risk adjustment literature. Schokkaert and Van de Voorde (2000) have argued that while there is a difference between the positive exercise of *explaining* medical care expenditure (or use) and the normative issue of justifying medical expenditure (or use) differences, the results of the former exercise have relevance for the second. Drawing on the theory of fair compensation, they show that failure to include 'responsibility variables' (which *do not* need to be compensated for in the capitation formula) in the equation used for estimating the effect of 'compensation variables' (which *do* need to be compensated for) may give rise to omitted variable bias in the determination of the 'appropriate' capitations (or fair compensations). Their proposed remedy to this problem is to include the 'omitted variables' in the estimation equation but to 'neutralize' their impact by setting these variables equal to their means in the need-prediction equation.

A similar argument to Schokkaert and Van de Voorde was made and taken further by Gravelle (2003) in the context of the measurement of income-related inequality of health or health care. He uses an 'augmented partial concentration index' which is defined as the (directly) standardised concentration index, but controlling for income and other non-standardising variables in the process. This can be obtained from the regression-based decomposition of the concentration index.

One important problem with measuring horizontal inequity and applying the decomposition analysis is that the dependent variable in health care demand models is typically specified as a nonlinear function of the regressors: for example, in van Doorslaer, Koolman, and Jones (2004) the empirical models of health care use are based on logistic and truncated and generalized negative binomial regression models, which are intrinsically nonlinear. So long as the model is linear, then the Schokkaert and Van de Voorde (2000) approach of estimating the linear regression and then neutralizing the non-need variables by setting them equal to their mean (or, in fact, any constant value) and the decomposition approach lead to the same measure of horizontal inequity (van Doorslaer, Koolman and Jones, 2004). This does not hold for a nonlinear model, as the linear decomposition does not apply.

To compute horizontal inequity in the context of a nonlinear model, again we have used a two-step approach. In the first step we predict need-expected utilisation based on the actual values of the x_n variables, but these predictions are contingent on the level of the non-need variables (x^r and x^p) that is selected. By analogy with the linear case, we have chosen to set the non-need variables equal to their sample means. So:

$$\hat{y}_i = E(y \mid x_i^n, \bar{x}^r, \bar{x}^p) = G\left(\sum_n \hat{\beta}_n x_i^n + \sum_r \hat{\beta}_r \bar{x}_i^r + \sum_p \hat{\beta}_p \bar{x}_i^p\right) \qquad (3)$$

In the second step the HI index is then obtained by subtracting the concentration index of $\hat{y}$ from the concentration index of y. A complication compared to the linear case is that the

HI index for the nonlinear model is contingent on the values used for the non-need variables and therefore their effect is not completely neutralised.

4. Econometric methods

4.1 The latent class hurdle model

To compute indices of horizontal equity that exploit the panel data dimension of the ECHP we estimate latent class hurdle models, a model developed by Bago d'Uva (2005). A latent class (or finite mixture) framework is adopted in which individual effects are approximated using a discrete distribution. This framework offers an alternative representation of heterogeneity, where individuals are drawn from a finite number of latent classes. The latent class framework has been used previously in models for health care utilisation with cross-sectional individual data. Deb and Trivedi (2002) note that this framework "provides a natural representation of the individuals in a finite number of latent classes, that can be regarded as types or groups". The segmentation can represent individual unobserved characteristics such as unmeasured health status not sufficiently accounted for by the observed measures. Deb and Trivedi (1997, 2002), Deb and Holmes (2000), and Jimenez-Martin et al (2002) estimate finite mixture models for count measures of health care use, in which a NB distribution is assumed within each latent class. Atella et al (2004) develop a latent class model for the joint decisions of consulting 3 types of physician. The authors assume that, within a latent class, each decision can be modelled by an independent probit, so the joint distribution of the 3 binary outcomes is a product of probits. Deb (2001) develops a random effects probit in which the distribution of the random intercept is approximated by a discrete density. This approximation relaxes the normality assumption for the distribution of the random effects. Deb (2001) applies this model to a cross-section of individuals clustered in families, where the random effect represents unobserved family effects. It is assumed therefore that all individuals in each family belong to the same latent class. This approach aims to approximate the distribution of the random (family) intercepts, whereas the responses to the explanatory variables are not allowed to vary across latent classes.

There are a number of applications of latent class models in other fields (e.g. Wang, 1998; Wedel, 1993; Nagin, 1993; Uebersax, 1999). Greene (2001) notes that most applications have not used panel data. However, according to this author, the latent class model is "only weakly identified at very best by a cross-section". Additionally, Greene notes that the richness of the panel in terms of cross-group variation improves the potential to estimate the model. The recent implementation of latent class models for panel data in LIMDEP 8.0 (Greene, 2002) suggests that this approach may become more popular in the near future. In the context of smoking behaviour, Clark and Etile (2003) use the latent class framework to approximate the continuous distribution of the individual effects in a dynamic random effects bivariate probit model. Clark et al (2005) develop a latent class ordered probit model for reported well-being, in which individual time invariant heterogeneity is allowed both in the intercept and in the income effect.

This paper uses a panel of individuals across time. Individuals i are observed T_i times, where T_i =1,… ,7. Let y_{it} represent the number of visits in year t. Denote the observations of the dependent variable over the panel as $y_i =[y_{i1} ,..., y_{iTi}]$. Consider that individual i belongs to a latent class j, j=1,…,C, and that individuals are heterogeneous across classes. Conditional on the covariates considered, there is homogeneity within a given class j. Given the class that individual i belongs to, the dependent variable in a given year t, y_{it}, has density $f_j(y_{it}| x_{it},\theta_j)$. The joint density of the dependent variable over the observed periods is a product of T_i independent densities $f_j(y_{it}| x_{it},\theta_j)$, given class j. The probability of belonging to class j is π_{ij}, where $0<\pi_{ij} <1$ and $\Sigma^{C}_{j=1}\pi_{ij} =1$. Unconditionally on the latent class the individual belongs to, the joint density of $y_i =[y_{i1} ,..., y_{iTi}]$ is given by:

$$g(y_i \mid x_i;\pi_{i1},...,\pi_{iC};\theta_1,...,\theta_C)=\sum_{j=1}^{C}\pi_{ij}\prod_{t=1}^{T_i} f_j(y_{it} \mid x_{it};\theta_j) \qquad (4)$$

where x_i is a vector of covariates, including a constant and θ_j are vectors of parameters.

Conditional on the class that the individual belongs to, the number of visits in period t, y_{it}, is assumed to be determined by a hurdle model. The underlying distribution for the two stages of the hurdle model is the NegBin. Formally, for each component $j = 1,\ldots C$, it is assumed that the probability of zero visits and the probability of observing y_{it} visits, given that y_{it} is positive, are given by the following expressions:

$$\begin{aligned}
f_j(0 \mid x_{it}; \beta_{j1}) &= P[y_{it} = 0 \mid x_{it}, \beta_{j1}] = \left(\lambda_{j1,it}^{1-k} + 1\right)^{-\lambda_{j1,it}} \\
f_j(y_{it} \mid x_{it}; \beta_{j2}) &= \frac{\Gamma\left(y_{it} + \dfrac{\lambda_{j2,it}^{k}}{\alpha_j}\right)\left(\alpha_j \lambda_{j2,it}^{1-k} + 1\right)^{-\frac{\lambda_{j2,it}^{k}}{\alpha_j}}\left(1 + \dfrac{\lambda_{j2,it}^{k-1}}{\alpha_j}\right)^{-y_{it}}}{\Gamma\left(\dfrac{\lambda_{j2,it}^{k}}{\alpha_j}\right)\Gamma(y_{it} + 1)\left[1 - \left(\alpha_j \lambda_{j2,it}^{1-k} + 1\right)^{-\frac{\lambda_{j2,it}^{k}}{\alpha_j}}\right]}
\end{aligned} \tag{5}$$

where $\lambda_{j1,it} = \exp(x'_{it}\,\beta_{j1})$, $\lambda_{j2,it} = \exp(x'_{it}\,\beta_{j2})$, α_j are overdispersion parameters and k is an arbitrary constant (most commonly set equal to 0 or 1, which corresponds to the NB1 and NB2 models, respectively).

Similarly to the hurdle model, the fact that β_{j1} can be different from β_{j2} reflects the possibility that the zeros and the positives are determined by two different decision processes. In other words, the determinants of care are allowed to affect differently the two stages of the decision process regarding the number of visits to the doctor: i) the probability of seeking care and ii) the number of visits, given that this is positive. On the other hand, having $[\beta_{j1}, \beta_{j2}] \neq [\beta_{l1}, \beta_{l1}]$ for $j \neq l$, reflects the differences between the latent classes. The same set of regressors is considered in both parts of the model. As to the variation between classes, it can be assumed that all the slopes are the same, varying only the constant terms, $\beta_{j1,0}$ and $\beta_{j2,0}$, and the overdispersion parameters α_j. This represents a case where there is unobserved individual heterogeneity but not in the responses to the covariates (as in the model used in Deb, 2001). The most flexible version allows α_j and all elements of β_{j1} and β_{j2} to vary across classes. The finite mixture hurdle model also accommodates a mixture of sub-populations for which health care use is determined by a NegBin model (the two decision processes are indisguishable) and sub-populations for which utilisations is determined by a hurdle model. This is obtained by setting $\beta_{j1} = \beta_{j2}$, for some classes.

The discrete distribution of the heterogeneity has C mass points. In previous empirical applications of latent class model to health care utilisation, class membership probabilities were taken as parameters $\pi_{ij}=\pi_j$, j=1,… ,C, to be estimated along with θ_1, …, θ_C (Deb and Trivedi, 1997 and 2002; Deb and Holmes, 2000; Deb, 2001; Jimenez et al; 2002, Atella et al, 2004). These can also be parameterised as functions of time invariant individual characteristics z_i. In this case, class membership is modelled as a multinomial logit (as in, for example, Clark and Etilé, 2003; Clark et al, 2005):

$$\pi_{ij} = \frac{\exp\left(z'_i \gamma_j\right)}{\sum_{g=1}^{C} \exp\left(z'_i \gamma_g\right)}, \quad j = 1,..., C, \tag{6}$$

with γ_C=0. This uncovers the determinants of class membership.[1] In a panel data context, this parameterisation provides a way to account for the possibility that the observed regressors may be correlated with the individual effect. Let $z_i = \bar{x}_i$ be the average over the observed panel of the observations on the covariates. This is in line with what has been done in recent studies to allow for the correlation between covariates and random effects, following the suggestion of authors such as Mundlak (1978). The vectors of parameters θ_1, …, θ_C, γ_1, …, γ_{C-1} are estimated jointly by maximum likelihood.

After the estimation of the model, it is possible to calculate the posterior probability that each individual belongs to a given class. The posterior probability of membership in class j is given by:

---

[1] In previous latent class models for health care utilisation, this has been done through posterior analysis.

$$P[i \in j] = \frac{\pi_{ij} \prod_{t=1}^{T_i} f_j\left(y_{it} \mid x_{it}; \theta_j\right)}{\sum_{j=1}^{C} \pi_{ij} \prod_{t=1}^{T_i} f_j\left(y_{it} \mid x_{it}; \theta_j\right)} \tag{7}$$

The individuals can then be assigned to the class with the highest posterior probability.

The latent class panel data model accounts for the panel feature of the data in a flexible way that assumes no distribution for the unobserved individual effects. It can also be seen as a discrete approximation of an underlying continuous mixing distribution (Heckman, 1984). The number of points of support needed for the finite mixture model is low, usually two or three. The specification used here allows for correlation between latent heterogeneity and the covariates. The conventional fixed effects models that have been developed for binary dependent variables (conditional logit) and for counts (fixed effects Poisson and NB) also offer a distribution-free approach to the individual heterogeneity that is robust to correlation between covariates and individual effects. However, although fixed effects models account for intercept heterogeneity, they do not accommodate different responses to the covariates across individuals, while the latent class model accommodates both intercept heterogeneity and slope heterogeneity. Furthermore, fixed effects models do not allow the estimation of the effects of time invariant regressors. In these models, the coefficients of time invariant regressors are absorbed into the individual effect and, thus, are not identified.

Estimation was done by maximum likelihood using TSP 4.5 (32). The Newton method is used for the models with one component. The Broyden-Fletcher-Goldfarb-Shanno quasi-Newton algorithm is implemented in TSP as an option of the maximum likelihood estimation. This method is used to estimate the latent class model. Due to the possibility of convergence to local maxima in mixture models, the estimation should be repeated using different sets of starting values for the parameters being estimated. These starting values can be obtained as combinations of the estimates of the one component version of the model. Moreover, estimates of restricted versions of the model (for example, with constant slopes, or with constant class membership) can be used as starting values in the estimation of more flexible versions.

4.2 Inequity indices

In order to calculate inequity indices, we predict health care utilisation fixing the non-need variables at the sample means. The number of visits is predicted using the estimated panel data latent classes model. However, the calculation of predictions in this highly non-linear model is not straightforward. In particular, it is necessary to define whether the individual unobserved heterogeneity represents need, non-need or both. As noted by Van Ourti (2004), unobserved individual heterogeneity may reflect need factors (such as unobserved health) as well as non-need factors (such as health care preferences).

The key assumption that we make is that the predictions vary only with need, i.e., the horizontal equity norm is that there is equal treatment for equal need. Different assumptions regarding the nature of the individual unobserved heterogeneity require different procedures to predict utilisation. The following paragraphs present three possible methods of calculating the predicted number of visits, conditional on actual values for the need variables, x_{it}^{N}, and sample averages for the non-need variables $\bar{x}_{t}^{NN}$. The first and second options can be applied when the class membership probabilities, π_{ij} in equation (4) are parameters, π_j, $j = 1,\ldots,C$. The third option concerns a latent class model in which the class membership probabilities are defined as functions of the covariates, as in equation (5).

Option 1

Consider the latent class model in equation (4) and the class membership probabilities as parameters $\pi_{ij} = \pi_j$, j=1,...,C. The number of visits can be predicted in the following way:

$$\hat{y}_{it} = \sum_{j}^{C} \pi_j E_j \left[y_{it} \mid x_{it}^{N}, \bar{x}_{t}^{NN} \right] \tag{8}$$

As long as two individuals match in terms of the need characteristics x_{it}^{N}, their predicted utilisation is the same, even if they belong to different latent classes. Thus, the individual heterogeneity is treated as *non-need*. This treatment of the unobserved time-invariant heterogeneity resembles the treatment of the random effects in the panel data hurdle model in Van Ourti (2004).

Option 2

Consider again that the class membership probabilities are parameters $\pi_{ij} = \pi_j$, j=1,...,C. After the estimation of the model, it is possible to calculate the posterior class membership probabilities using equation (6). Then, each individual can be assigned to the class of highest posterior probability, $j*$. Finally, the predicted number of visits in year t can be calculated as the expected value of y_{it}, conditional on class $j*$, x_{it}^{N} and $\bar{x}_{t}^{NN}$:

$$\hat{y}_{it} = E_{j*}\left[y_{it} \mid x_{it}^{N}, \bar{x}_{t}^{NN}\right] \tag{9}$$

Given $j*$, the predictions vary only with need. Even amongst individuals that have the same values for the need variables, there is still variation in the predicted use to the extent that individuals belong to different latent classes. Therefore, the individual heterogeneity (represented by membership to different latent classes) is treated as *need*.

Option 3

Consider now that the class membership probabilities are specified as functions of the covariates as in equation (5). Similarly to the x's, the time-invariant determinants of class membership, z_i, can include both *need* and *non-need* factors. Thus, we are not restricting the individual unobserved heterogeneity to be solely need (non-need) as in the above options. The predictions can be computed as:

$$\hat{y}_{it} = \sum_{j}^{C} \hat{\pi}_j\left(z_i^{N}, \bar{z}^{NN}\right) E_j\left[y_{it} \mid x_{it}^{N}, \bar{x}_{t}^{NN}\right] \tag{10}$$

Since the class membership probabilities are computed for fixed values of the non-need variables, we are assuming that, across individuals, only the variation in π_{ij} (.) that is related to need is legitimate. All the individuals with the same need are attributed the same class membership probabilities. Similarly, conditional on the latent class, the predictions vary only according to need. Therefore, the resulting predictions, unconditional on the latent class, $\hat{y}_{it}$, vary only with the observed need factors. However, there is still some unexplained variation since we predict the probabilities of class membership, not the actual classes to which the individual belongs. In this setting, that remaining variation is treated as non-need.

5 Results

5.1 Income-related inequality in doctor visits by country and wave

GP visits

Figures 1 to 13 show the concentration curves of the number of GP visits for each of the 13 EU countries included in the analysis. Each of the charts includes the concentration curves for all the waves available for the respective country as well as the long-run concentration curve. In general, the curves lie above the 45º line (except for some which lie slightly below the equality line in the bottom of the income distribution), showing a pro-poor distribution of GP visits. For most of the countries, the variation in the curves across time appears to occur especially in the bottom of the income distribution (showing that the distribution of GP visits among the poor varies more across time than the distribution of GP visits among the rich). Exceptions are found in the case of Italy (Figure 8), that presents little variation across time, Denmark and Greece, for which variation is observed throughout the curves, and Germany, which curves vary less in the bottom of the income distribution. It does not appear to be possible to identify patterns in the evolution of the concentration curves across time.

The countries that present larger pro-poor inequalities in the number of GP visits (concentration curves further above the equality line) are Belgium (Figure 2), Greece (Figure 6) and Ireland (Figure 7). For Greece, it is possible to distinguish 2 sets of curves (the curves of waves 3, 5, 6 and 8 tend to be below the others). Finland is the country which curves are nearer to the equality line, showing the smallest degree of pro-poor inequality, throughout the panel.

The largest variations across time are observed for Denmark, Ireland, and Greece. While the curves for Greece maintain a similar shape, this is not the case for Denmark and Ireland, which curves vary not only in the distance from the 45º line but also in shape. The countries that show less variation in the curves are the United Kingdom (Figure 13, 2 waves only) and Italy (Figure 8).

The information on income-related inequality in GP visits contained in Figures 1 to 13 is summarised in the short-run concentration indices, CI^t, and the long-run concentration indices, CI^T, in Table 4. In all the countries, across waves as well as in the long-run, all the concentration indices are negative, reflecting the pro-poor inequality already identified in the analysis of the concentration curves. Within country, the indices vary across time, but they do not follow any recognisable pattern. For Finland, all the short-run indices (on average, -0,024) as well as the long-run index (-0.33) are smaller (in absolute value) than for the remaining countries, which is in line with the fact that the concentration curves for this country are the nearest to the equality line (as seen above). The largest (in absolute value) short-run concentration indices are observed for Belgium and Ireland in all the waves, except for waves 2 and 4, when pro-poor inequality was larger for Greece. Averaged across waves, short-run inequality is larger (in absolute value) for Ireland (-0.144), followed by Belgium (-0.143) and Greece (-0.133). In the long-run, pro-poor inequality is larger for Greece (-0.166), Ireland (-0.159) and Belgium (-0.158).

The mobility indices in Table 4, MI^T, give the discrepancy between the short-run and the long-run measures of income-related inequality, for each country. These indices are negative for all the countries, except for Luxembourg (0.057), the Netherlands (0.037) and Germany (0.011). This means that, for 10 of the countries included in the analysis, the short-run measure understates long-run pro-poor inequality between 7% (UK) and 40% (Finland).

Specialist visits

The concentration curves of specialist visits (by wave and long-run) for each of the 13 EU countries are shown in Figures 14 to 26. For all the countries, the curves present greater variation across time than in the case of GP visits seen above. The variation across countries is close to what is observed for GP visits, however, the location of the curves is very different. In Finland, Italy and Portugal (Figures 17, 21 and 24), the curves lie always below the equality line, showing pro-rich inequality in the distribution of specialist visits, in all the waves. In Belgium, Greece and the Netherlands (Figures 14, 19 and 23), the concentration curves show pro-poor inequality in all the waves, which is of similar magnitude in these three countries. In the remaining countries, the curves are above the 45º line in some waves and below in others, or cross that line in some waves.

Denmark, Ireland and Luxembourg (Figures 16, 20 and 22) are the countries for which the curves vary the most across waves. For Luxembourg, the curve for wave 3 is above the 45º line, whereas the curve for wave 2 crosses that line, being near the line throughout the income distribution. In the case of Ireland, the curve for wave 8 lies above the equality line, while the curves of other waves are mainly below the equality line. The long-run concentration curve is near the 45º line, mainly below this one. For Denmark, the curve for wave 8 lies above the equality line and mainly above the curves of the other waves. The long-run curve for Denmark lies very close to the equality line. Throughout the observed waves, the concentration curves of Germany, Spain and the UK (figures 18, 25 and 26) are consistently near the 45º line.

The information contained in the concentration curves in Figures 14 to 26 is summarised using concentration indices, CI^t, and long-run concentration indices, CI^T, in Figure 5. Throughout the observed period, the largest short-run indices are observed for Portugal (on average, 0.117) and Finland (on average, 0.091), except for wave 6, when the index for Italy is slightly larger than the one of Finland. The indices for Portugal are always greater than the ones for the remaining countries, except for Finland in wave 3 (same as Portugal) and wave 4 (above Portugal). The long-run indices are also larger for Portugal (0.110) and Finland (0.102). The short-run indices for Greece (on average, -0.052), the Netherlands (on average, -0.041), Belgium (on average, -0.040) and the United Kingdom (on average, -0.017), are always negative. The indices for these four countries are the lowest in all the waves, except for wave 2, when the index for Denmark is lower than the ones of the Netherlands and the United Kingdom, wave 3, when Germany has the lowest index, and wave 8, when Ireland and Denmark have the lowest indices. In the long-run, the largest pro-poor inequalities in the number visits to a specialist are registered for Greece (-0.076), followed by Belgium (-0.038).

The proportional difference between the long-run and the weighted average of short-run indices is given by the mobility indices in Table 5, MI^T. For Belgium, Luxembourg and the Netherlands, the positive mobility indices, coupled with negative short-run concentration indices (on average) and negative long-run concentration indices, show that the long-run pro-poor inequality is smaller than the short-run pro-poor inequality. The negative mobility indices for Greece and the United Kingdom mean that the pro-poor inequality observed in these countries is larger in the long-run than in the short-run. For Spain, Italy and Portugal, the mobility indices as well as the short-run concentration indices (on average) and the long-run concentration indices, are positive, showing that the long-run pro-rich inequality is smaller than the short-run pro-rich inequality. Finally, amongst the countries that show pro-rich inequality in the short-run (on average) and in the long-run, long-run inequality is understated by the short-run measure in Germany, Ireland, Austria and Finland.

5.2 Cross-country comparison of long-run income-related inequality in doctor visits

We look now into more detail at the long-run income-related inequality in doctor visits. In particular, we compare the long-run concentration curves across countries. We start by plotting the curves of all the countries in the same graph (Figure 28 for GP visits and figure 31 for specialist visits) in order to identify any dominance relationships that might exist. When the concentration curve for one country lies everywhere above or below another, the ranking of those two countries by degree of inequality is unambiguous. When the concentration curves of a pair of countries cross, the comparison of the inequality levels requires a summary measure such as the concentration index.

GP visits

Figure 27 presents the long-run concentration curves of the 13 countries considered in the analysis. The relative positions of the curves are not clear from this figure. Nevertheless, it possible to observe that the curve of Finland is everywhere below all the others, while the curves of Greece, Belgium, and Ireland tend to be above. The following figures plot selections of curves, with the aim of better identifying their relative positions. Figure 28 includes the curves of Greece, Belgium and Ireland as well as the curves of Spain and Luxembourg. It is shown that, although the curves of the former three countries are almost everywhere above the latter ones, they cross the curve of Luxembourg in the extremes and the curve of Spain at about the 8^{th} income decile. It is also possible to note that the curves of Belgium and Ireland almost overlap, and cross the curve of Greece at the 2^{nd} and the 6^{th} deciles (with Greece below between these two points). In Figure 29, the curve of Greece is plotted together with all the curves that are everywhere below it. Similarly, Figure 30 plots the curves of Ireland and all the others that are everywhere below that one. It is therefore clear that pro-poor long-run income-related inequality in GP visits is larger for Greece, Ireland and Belgium than for the remaining countries. However, the curves are not sufficient to rank those three countries as to the level of inequality. This can be done by means of the concentration indices presented in Table 7 in increasing order. The three indices do not differ much, being the lowest observed for Greece (-0.166, representing the highest pro-poor inequality), while the indices for Ireland and Belgium are almost equal (-0.159, -0.158). The indices also confirm that the country with the lowest long-run income-related inequality of GP visits is Finland, as it was clear from the analysis of the concentration curves. Moreover, as the analysis in the previous section has already shown, this also holds for the short-run indices across all waves.

Specialist visits

The long-run concentration curves of specialist visits for all the analysed countries are plotted in Figure 31. As it was notes above in the analysis by country and wave, inequality in specialist visits is pro-rich in some cases and pro-poor in other cases. Figure 32 shows that the long-run concentration curves of Finland, Portugal, Italy, Spain and Austria are everywhere below the equality line, meaning that there is pro-poor long-run inequality in specialist visits in these countries. The long-run curves for Finland and Portugal lie below the ones of the remaining countries throughout the income distribution. The concentration curves of these two countries almost overlap in the first half, being Portugal below in the second half. The information provided by the concentration curves is summarised in the concentration indices in Table 6. These make clear that Portugal has the highest level of pro-rich inequality (index equals 0.110), followed closely by Finland (0.102), and at a greater distance, by Italy and Austria (both 0.042). The concentration curves for Germany and Ireland cross the 45º line, but are mainly below this one (see Figure 33), which corresponds to positive concentration indices (0.013 and 0.027, respectively). Belgium, Greece, the Netherlands and Luxembourg have pro-rich inequality, as shown by the concentration curves in Figure 34. These are everywhere above the equality line, except for the curve of Belgium that is slightly below in the bottom of the income distribution. The curve for Greece is everywhere above the ones of the remaining countries, except for the bottom of the income distribution where Luxembourg is slightly above. Greece is therefore the country that presents the largest pro-poor long-run inequality, which is confirmed by the long-run concentration indices in Table 6 (-0.076 for Greece, followed by Belgium, -0.036). In the case of Denmark and the United Kingdom, the curves cross the 45º line in more than one point (see Figure 35). It can however be noted that these curves lie near the equality line, mainly above this one. Accordingly, the concentration indices for those two countries are negative and small in absolute value (-0.018 for the United Kingdom, and -0.011, for Germany).

5.3 Preliminary results of econometric models

We estimate cross-sectional hurdle and panel data LC hurdle models for specialist and GP visits. The standard hurdle model corresponds to a (degenerate) LC model with only one component, in which the panel structure of the data is not accounted for. At a first stage of the econometric analysis, the LC hurdle model is defined with constant class membership probabilities, $\pi_{ij} = \pi_j$, j=1,...,C, and with 2 latent classes, C=2 (equation(4)). The underlying distribution in both stages, for all the models, is a NB2, i.e. k=1 in equation (5). We present the estimated effects of income in the hurdle model for all the countries covered by the analysis. For each type of doctor, we compare those results with the preliminary ones obtained in the LC hurdle, for two countries.

GP visits

Table 8 presents the estimated effects of income in hurdle models for GP visits, conditional on the remaining need and non-need factors considered. The estimated income effects on the probability of visiting a GP are positive for most of the countries. These positive effects are significant for Denmark, Finland, Ireland, The Netherlands and Portugal. Germany and Spain show significantly negative income effects on the first part of the model. The second part of the model shows very different results. These are negative for 10 countries, being significant for 8 of these. The estimated income effect on the conditional number of visits is only significantly positive in the case of Austria. Only three countries exhibit the income effects of the same sign in both parts: for Denmark, they are both positive (significant in the first part); for Luxembourg, they are both negative (significant in second part) and, for Spain, they are significantly negative in both parts. For 8 countries, the estimated effects are significant in only one part of the model: negative in the first part for Italy, Luxembourg, Belgium, Greece; negative in the second part for Germany; positive in the first part for Denmark and positive in the second part for Austria. The model estimated for the UK shows no evidence of income effects.

A comparison between the income effects obtained with the hurdle model and preliminary results given by the LC hurdle is given in Table 9, for Austria and Denmark. The two models are also compared as to maximised log-likelihood and Schwarz information criterion (BIC). It is shown that the panel data LC hurdle provides a considerable improvement in fit. This model outperforms the cross-section hurdle, even when the additional number of parameters is penalised for by the BIC. These results give support to the existence of unobserved time-invariant individual heterogeneity, in the two examples analysed here. The LC hurdle identifies two latent classes of users. It is estimated that the class of low users represents 64% of population, in the case of Austria, and 54%, in the case Denmark. For Austria, the hurdle model estimates a positive and significant effect of income in the second part of the model. The LC hurdle allows for the effects to be different across latent classes. This leads to a positive and significant income effect on the second stage, only for the high users of primary care. The remaining effects are negative and insignificant. For Denmark, the hurdle model results in a significantly positive income effect in the probability of visiting a GP. The LC hurdle further identifies that this effect is significant only for low users.

Specialist visits

The estimated effects of income in the cross-sectional hurdle model are shown in Table 10. It is noticeable that the income effects are mostly positive. In particular, the effects of income on the probability of seeking specialist care are always positive and significant, with the largest values being observed for Portugal and Finland, followed by Ireland and Denmark. The income effects on the expected number of specialist visits, given that there is at least one visit, are mostly positive, yet not as significant as in the first part of the model. These are positive and significant for Austria, Germany, Greece and Portugal, and insignificant for 8 countries (5 positive and 3 negative). The income effect on the second part of the model is significantly negative only for Luxembourg. Portugal and Finland show the largest evidence of a positive income effect on specialist visits, conditional on need and remaining non-need factors considered. These are also the countries that present the largest long-run income-related inequality, not controlling for other need and non-need factors (Table 6).

Table 11 compares the preliminary estimation results of the income effects in the LC hurdle model with the ones obtained with the hurdle model, for Finland and Ireland. For both cases, accounting for the panel structure of the data by means of the LC hurdle leads to a considerable improvement in fit. The Schwarz information criterion (BIC) favours the LC hurdle over the hurdle model, even penalising for the inclusion of additional parameters. Moreover, the LC hurdle identifies two latent classes of users, with different income effects in the two parts of the decision process. In the case of Finland, the hurdle model estimates a positive and significant impact of income on the probability of visiting a specialist and a positive and insignificant effect on the conditional number of visits. The LC hurdle model shows further evidence that the positive income effect observed in the first stage is greater for low users of specialist care, which are estimated to represent 65% of the population. The effects of income that are obtained with the hurdle model for Ireland exhibit a similar pattern to what is obtained for Finland. The LC hurdle model however shows that, for the low users of care (72%), income impacts positively and significantly both stages, whilst, for high users, that is only the case of the first stage.

6. Preliminary conclusions

This paper presents results on income-related inequalities in health care (GPs and specialists), prior to need adjustment. It is shown that there is pro-poor inequality in the number of GP visits across countries and waves. The results regarding inequality in specialist visits are substantially different, showing mostly pro-rich inequalities. In particular, we observe pro-rich long-run inequality in specialist visits for 7 of the 13 countries considered here. The analysis of long-run inequalities confirms some of the results on short-run inequalities presented in van Doorslaer, Koolman & Jones (2004). Portugal shows the highest long-run pro-rich inequality in specialist visits. Ireland and Greece show the highest long-run pro-poor inequality in GP visits. There are however some new findings. Finland (that was not included in van Doorslaer, Koolman & Jones) presents the second highest level of pro-rich long-run inequality in specialist visits. Moreover, this country has the lowest pro-poor long-run inequality in GP visits. In the previous study, Austria presented the lowest pro-poor short-run inequality in GP visits, which is not confirmed by the long-run measure. Ireland is not one of the countries with higher pro-rich long-run inequality in specialist visits, despite having been second only to Portugal in van Doorslaer, Koolman & Jones (2004).

The next step of this work will be to complete the estimation of latent class panel data hurdle models for GP and specialists visits. The estimation results will then be used for the computation of predictions of need-expected health care use according to the three options discussed here. Each of these options implies a different assumption regarding the nature of individual unobserved heterogeneity. Comparison of the resulting horizontal inequity indices will allow us to assess the extent to which those assumptions affect the measurement of horizontal inequity. Ultimately, we will examine whether panel data gives a different picture of inequalities and inequities in health care use across European countries than what is given by cross-sectional data.

References

Atella, V., Brindisi, F., Deb, P., Rosati, F.C. Determinants of access to physician services in Italy: a latent class seemingly unrelated probit approach. *Health Economics* 2004; 13: 657-668.

Bago d'Uva, T., Latent class models for health care utilisation, HEDG Working Paper 05/01, 2005, University of York.

Clark, A., Etilé, F. Don't give up on me baby: Spousal correlation in smoking behaviour. *Delta Working Papers* 2003; 25. Delta École Normale Supérieure.

Clark, A., Etilé, F., Postel-Vinay, F., Senik, C., Van der Straeten, K. Heterogeneity in reported well-being: Evidence from twelve European countries. *The Economic Journal* 2005; 115: C118-C132.

Deb, P. A discrete random effects probit model with application to the demand for preventive care. *Health Economics* 2001; 10: 371-383.

Deb, P., Holmes, A.M. Estimates of use and costs of behavioural health care: A comparison of standard and finite mixture models. *Health Economics* 2000; 9: 475-489.

Deb, P., Trivedi, P.K. Demand for medical care by the elderly: a finite mixture approach. *Journal of Applied Econometrics* 1997; 12: 313-336.

Deb, P., Trivedi, P.K. The structure of demand for health care: latent class versus two-part models. *Journal of Health Economics* 2002; 21: 601-625.

Gravelle, H. (2003) Measuring income related inequality in health: standardization and the partial concentration index. *Health Economics* 2003; 12: 803-819.

Greene, W., Fixed and Random Effects in Nonlinear Models. Unpublished manuscript, New York University, Department of Economics, Stern School of Economics, 2001.

Greene, W., LIMDEP, Version 8.0, Econometric Software, Plainview, New York, 2002.

Heckman, J., Singer, B. A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica* 1984; 52: 271-320.

Jimenez-Martin, S., Labeaga, J.M., Martinez-Granado, M. Latent class versus two-part models in the demand for physician services across the European Union. *Health Economics* 2002; 11: 301-321.

Jones, A.M. and A. López Nicolás 'Measurement and explanation of socioeconomic inequality in health with longitudinal data', *Health Economics* 2004; 13, 1015-1030.

Kakwani, N.C., A. Wagstaff, and E. Van Doorslaer, Socioeconomic inequalities in health: Measurement, computation and statistical inference. *Journal of Econometrics*, 1997. **77**(1): p. 87-104.

Kakwani, N.C., *Income Inequality and Poverty: Methods of Estimation and Policy Applications*. 1980, New York: Oxford University Press.

Lambert, P., *The distribution and redistribution of income: A mathematical analysis*. 2nd ed. 1993, Manchester: Manchester University Press.

Mundlak Y. On the pooling of time series and cross-section data. *Econometrica* 1978; 46: 69-85.

Nagin, D., Land, K. Age, Criminal Careers, and Population Heterogeneity: Specif{}ication and Estimation of a Nonparametric Mixed Poisson Model. *Criminology* 1993; 31:327-362.

Riphahn, R., Wambach, A., Million, A. Incentive effects in the demand for health care: a bivariate panel count data estimation. Journal of Applied Econometrics 2003. 18: 387-405.

Schokkaert, E., van de Voorde, C. (2000). *Risk selection and the specification of the risk adjustment formula*. Centre for Economic Studies Discussion Paper, 2000/011, University of Leuven, Leuven,

Uebersax, J.S. Probit Latent Class Analysis With Dichotomous or Ordered Category Measures: Conditional Independence/Dependence Models. *Applied Psychological Measurement* 1999; 23: 283-297.

van Doorslaer, E., X. Koolman, and F. Puffer, *Equity in the use of physician visits in OECD countries: has equal treatment for equal need been achieved?*, in *Measuring Up: Improving Health Systems Performance in OECD Countries*, OECD, Editor. 2002, OECD: Paris. 225-248.

van Doorslaer, E., Koolman, X. Jones, A.M. (2004) "Explaining income-related inequalities in doctor utilisation in Europe", *Health Economics* 2003; .

Van Ourti, T. Measuring horizontal inequity in Belgian health care using a Gaussian random effects two part count data model, *Health Economics* 2004; 13: 705-724.

Wagstaff, A., van Doorslaer, E. (2000). Measuring and testing for inequity in the delivery of health care. *Journal of Human Resources*, **35**(4): 716-733.

Wagstaff, A., P. Paci, and E. van Doorslaer, On the measurement of inequalities in health. *Social Science and Medicine*, 1991. **33**: p. 545-557.

Wang, P., Cockburn, I.M., Puterman, M.L. Analysis of Patent Data - A Mixed Poisson Regression Model Approach. *Journal of Business and Economic Statistics* 1998;16: 27-41.

Wedel, M., DeSarbo, W.S., Bult, J.R., Ramaswamy, V. A Latent Class Poisson Regression Model for Heterogeneous Count Data. *Journal of Applied Econometrics* 1993; 8: 397-411

Table 1: Average equivalised annual household income (real terms, common currency)[2]

	95	96	97	98	99	00	01
Austria	14355	14216	13765	13727	14133	14325	14003
Belgium	14280	14330	14240	14513	15318	15533	15496
Denmark	13632	11365	13641	13939	14115	14057	14071
Finland		10946	11046	11340	11726	11777	12090
Germany	14234	14441					
Greece	7273	7368	7655	8168	8325	8536	8663
Ireland	11085	11476	12091	13046	13559	13344	14505
Italy	10088	10111	10078	10589	10935	11098	11170
Luxembourg	22242	22812					
Netherlands	12526	12709	12688	12754	13141	13033	13145
Portugal	7478	7717	8009	8150	8530	8772	9273
Spain	8537	8692	8827	9223	9777	10405	10952
UK	13469	14040					

Table 2: Average number of specialist visits in previous year

	95	96	97	98	99	00	01
Austria	1.83	2.81	2.11	2.10	2.12	2.12	2.15
Belgium	1.85	1.82	1.83	1.97	1.93	1.87	1.95
Denmark	0.89	1.01	1.04	1.06	1.03	0.99	1.10
Finland		0.95	1.00	1.01	1.05	0.98	1.00
Germany	3.18	3.27					
Greece	1.63	1.69	1.98	1.59	1.74	1.80	1.78
Ireland	0.71	0.61	0.69	0.65	0.60	0.60	0.63
Italy	1.12	1.27	1.43	1.37	1.38	1.39	1.42
Luxembourg	2.02	2.30					
Netherlands	1.76	1.69	1.56	1.68	1.65	1.74	1.70
Portugal	1.18	1.29	1.38	1.28	1.34	1.40	1.28
Spain	1.73	1.51	1.67	1.63	1.58	1.55	1.71
UK	1.07	1.09					

[2] The values for Italy were multiplied by 1000.

Table 3: Average number of GP visits in previous year

	95	96	97	98	99	00	01
Austria	4.85	5.14	4.38	4.75	4.40	4.50	4.71
Belgium	5.18	5.06	4.96	5.13	5.10	5.05	4.92
Denmark	2.89	2.93	2.88	3.05	2.73	2.66	2.86
Finland		2.33	2.24	2.18	2.15	2.16	2.10
Germany	5.09	5.08					
Greece	2.15	2.24	2.31	1.93	1.98	2.11	1.87
Ireland	3.56	3.48	3.47	3.56	3.48	3.45	3.58
Italy	4.24	4.45	4.77	4.68	4.93	4.85	4.84
Luxembourg	3.15	3.19					
Netherlands	2.87	2.76	2.81	2.85	2.82	2.91	2.80
Portugal	3.26	3.39	3.33	3.35	3.50	3.28	3.07
Spain	3.78	3.57	4.24	3.74	3.54	3.45	3.87
UK	3.66	3.66					

Table 4: Short-run (CI^t) and long-run concentration indices (CI^T), and mobility indices (MI^T) for number of GP visits

Country	95	96	97	98	99	00	01	$\Sigma_\tau CI^t$	CI^T	MI^T
Austria	-0.090	-0.060	-0.087	-0.081	-0.079	-0.071	-0.084	-0.079	-0.090	-0.147
Belgium	-0.151	-0.138	-0.139	-0.135	-0.140	-0.148	-0.153	-0.143	-0.158	-0.102
Denmark	-0.074	-0.091	-0.086	-0.079	-0.109	-0.103	-0.123	-0.095	-0.104	-0.100
Finland		-0.021	-0.028	-0.040	-0.023	-0.008	-0.023	-0.024	-0.033	-0.395
Germany	-0.070	-0.093						-0.082	-0.081	0.011
Greece	-0.165	-0.113	-0.146	-0.126	-0.116	-0.148	-0.115	-0.133	-0.166	-0.244
Ireland	-0.131	-0.132	-0.144	-0.144	-0.152	-0.160	-0.147	-0.144	-0.159	-0.104
Italy	-0.053	-0.075	-0.080	-0.079	-0.063	-0.061	-0.060	-0.067	-0.076	-0.131
Luxembourg	-0.059	-0.104						-0.082	-0.077	0.057
Netherlands	-0.067	-0.056	-0.060	-0.074	-0.085	-0.097	-0.072	-0.073	-0.070	0.037
Portugal	-0.075	-0.065	-0.079	-0.091	-0.078	-0.072	-0.080	-0.077	-0.092	-0.199
Spain	-0.101	-0.085	-0.089	-0.085	-0.087	-0.113	-0.123	-0.098	-0.109	-0.122
UK	-0.104	-0.111						-0.108	-0.115	-0.070

Table 5: Short-run (CI^t) and long-run concentration indices (CI^T), and mobility indices (MI^T) for number of specialist visits

Country	95	96	97	98	99	00	01	$\Sigma_\tau CI^t$	CI^T	MI^T
Austria	0.033	0.034	0.051	0.066	0.023	0.025	0.012	0.035	0.042	-0.224
Belgium	-0.045	-0.037	-0.039	-0.060	-0.020	-0.038	-0.031	-0.039	-0.038	0.030
Denmark	-0.030	0.000	-0.002	0.031	-0.010	0.013	-0.116	-0.017	-0.011	0.376
Finland		0.112	0.116	0.072	0.072	0.111	0.064	0.091	0.102	-0.129
Germany	0.024	-0.009						0.008	0.013	-0.672
Greece	-0.032	-0.024	-0.059	-0.074	-0.055	-0.075	-0.046	-0.052	-0.076	-0.452
Ireland	0.060	0.078	-0.012	0.070	0.070	0.003	-0.119	0.021	0.027	-0.251
Italy	0.045	0.018	0.018	0.055	0.075	0.070	0.058	0.049	0.042	0.130
Luxembourg	0.007	-0.082						-0.040	-0.029	0.272
Netherlands	-0.017	-0.042	-0.056	-0.037	-0.037	-0.057	-0.040	-0.041	-0.025	0.388
Portugal	0.123	0.111	0.095	0.100	0.099	0.143	0.146	0.117	0.110	0.061
Spain	0.024	0.036	0.038	0.036	0.010	-0.025	0.000	0.017	0.015	0.117
UK	-0.009	-0.025						-0.017	-0.018	-0.056

Table 6: Long-run concentration indices (CI^T) of specialist visits (increasing order)

Country (waves)	CI^T
Greece (2-8)	-0.076
Belgium (2-8)	-0.038
Luxembourg (2-3)	-0.029
Netherlands (2-8)	-0.025
UK (2-3)	-0.018
Denmark (2-8)	-0.011
Germany (2-3)	0.013
Spain (2-8)	0.015
Ireland (2-8)	0.027
Italy (2-8)	0.042
Austria (2-8)	0.042
Finland (3-8)	0.102
Portugal (2-8)	0.110

Table 7: Long-run concentration indices (CI^T) of GP visits (increasing order)

Country (waves)	CI^T
Greece (2-8)	-0.166
Ireland (2-8)	-0.159
Belgium (2-8)	-0.158
UK (2-3)	-0.115
Spain (2-8)	-0.109
Denmark (2-8)	-0.104
Portugal (2-8)	-0.092
Austria (2-8)	-0.090
Germany (2-3)	-0.081
Luxembourg (2-3)	-0.077
Italy (2-8)	-0.076
Netherlands (2-8)	-0.070
Finland (3-8)	-0.033

Table 8: Estimated income effects in Hurdle models GP visits

Country	Income effects			
	P[Y>0]		E[Y|Y>0]	
Austria	-0.032	(-1.010)	0.045	(3.280)
Belgium	0.031	(1.040)	-0.063	(-5.320)
Denmark	0.130	(3.550)	0.018	(0.720)
Finland	0.085	(2.830)	-0.014	(-0.670)
Germany	-0.086	(-2.320)	0.005	(0.230)
Greece	0.022	(1.470)	-0.031	(-3.150)
Ireland	0.067	(2.440)	-0.097	(-6.040)
Italy	-0.022	(1.580)	-0.050	(-7.290)
Luxembourg	-0.098	(-1.100)	-0.118	(-2.980)
Netherlands	0.046	(2.300)	-0.091	(-6.780)
Portugal	0.106	(7.580)	-0.035	(-5.100)
Spain	-0.056	(-2.650)	-0.059	(-6.920)
UK	0.126	(0.180)	-0.015	(-0.680)

Note: t-statistics in parentheses

Table 9: Income effects in hurdle and latent class hurdle models for GP visits

Country	Model	Income effects P[Y>0]		Income effects E[Y|Y>0]		LogL	BIC
Austria	Hurdle						
		-0.032	(-1.010)	0.045	(3.280)	-84408.1	169620
	Latent Class Hurdle						
	Low users (64%)	-0.044	(-1.168)	-0.007	(-0.424)	-81486.2	164590
	High users (36%)	-0.132	(-0.992)	0.038	(2.013)		
Denmark	Hurdle						
		0.130	(3.550)	0.018	(0.720)	-58698.3	118165
	Latent Class Hurdle						
	Low users (54%)	0.200	(2.259)	0.023	(0.804)	-56895.4	115339
	High users (44%)	0.042	(0.720)	0.019	(0.286)		

Note: t-statistics in parentheses

Table 10: Estimated income effects in Hurdle models specialist visits

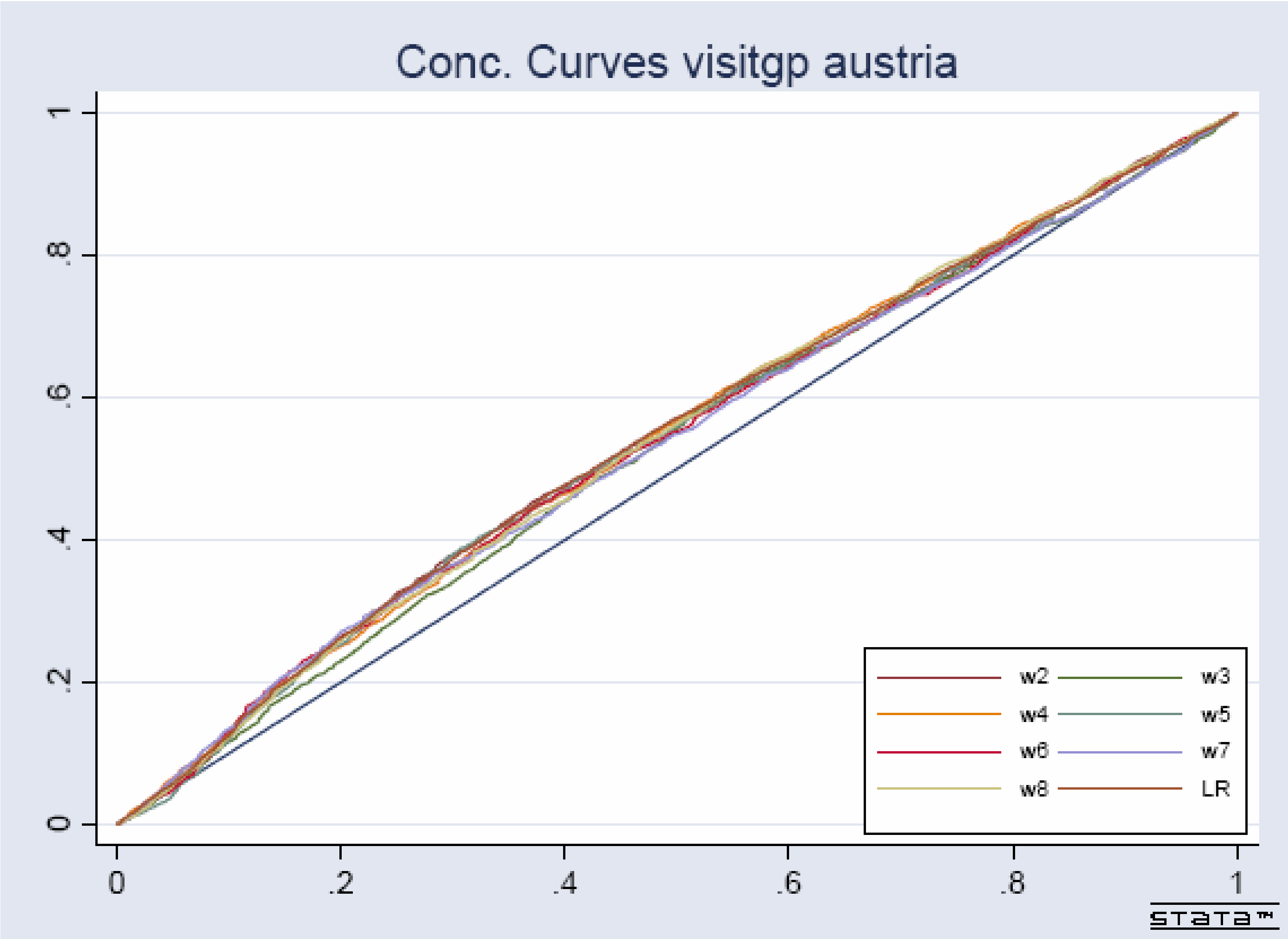
Country	Income effects P[Y>0]		Income effects E[Y|Y>0]	
Austria	0.285	(11.050)	0.139	(5.860)
Belgium	0.175	(7.700)	-0.015	(-0.660)
Denmark	0.298	(8.360)	0.036	(0.620)
Finland	0.462	(14.570)	0.019	(0.480)
Germany	0.186	(5.920)	0.159	(5.010)
Greece	0.202	(13.470)	0.048	(3.680)
Ireland	0.318	(10.810)	0.012	(0.300)
Italy	0.214	(17.350)	0.011	(0.770)
Luxembourg	0.276	(3.570)	-0.150	(-2.070)
Netherlands	0.118	(6.010)	-0.037	(-1.650)
Portugal	0.480	(32.580)	0.072	(5.180)
Spain	0.167	(13.990)	0.021	(1.600)
UK	0.289	(7.420)	-0.015	(-0.310)

Note: t-statistics in parentheses

Table 11: Comparison of income effects in hurdle and latent class hurdle models for specialists visits

		Income effects P[Y>0]		Income effects E[Y|Y>0]		LogL	BIC
Finland	Hurdle						
		0.462	(14.570)	0.019	(0.480)	-35664.4	72138
	Latent Class Hurdle						
	Low users (65%)	0.623	(9.463)	-0.007	(0.846)	-34479.8	70588
	High users (35%)	0.329	(4.705)	0.038	(0.622)		
Ireland	Hurdle						
		0.318	(10.810)	0.012	(0.300)	-31066.8	62938
	Latent Class Hurdle						
	Low users (72%)	0.245	(4.395)	0.232	(2.607)	-29990.9	61602
	High users (28%)	0.408	(6.450)	-0.066	(-1.284)		

Note: t-statistics in parentheses


Conc. Curves visitgp austria
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™
Figure 1

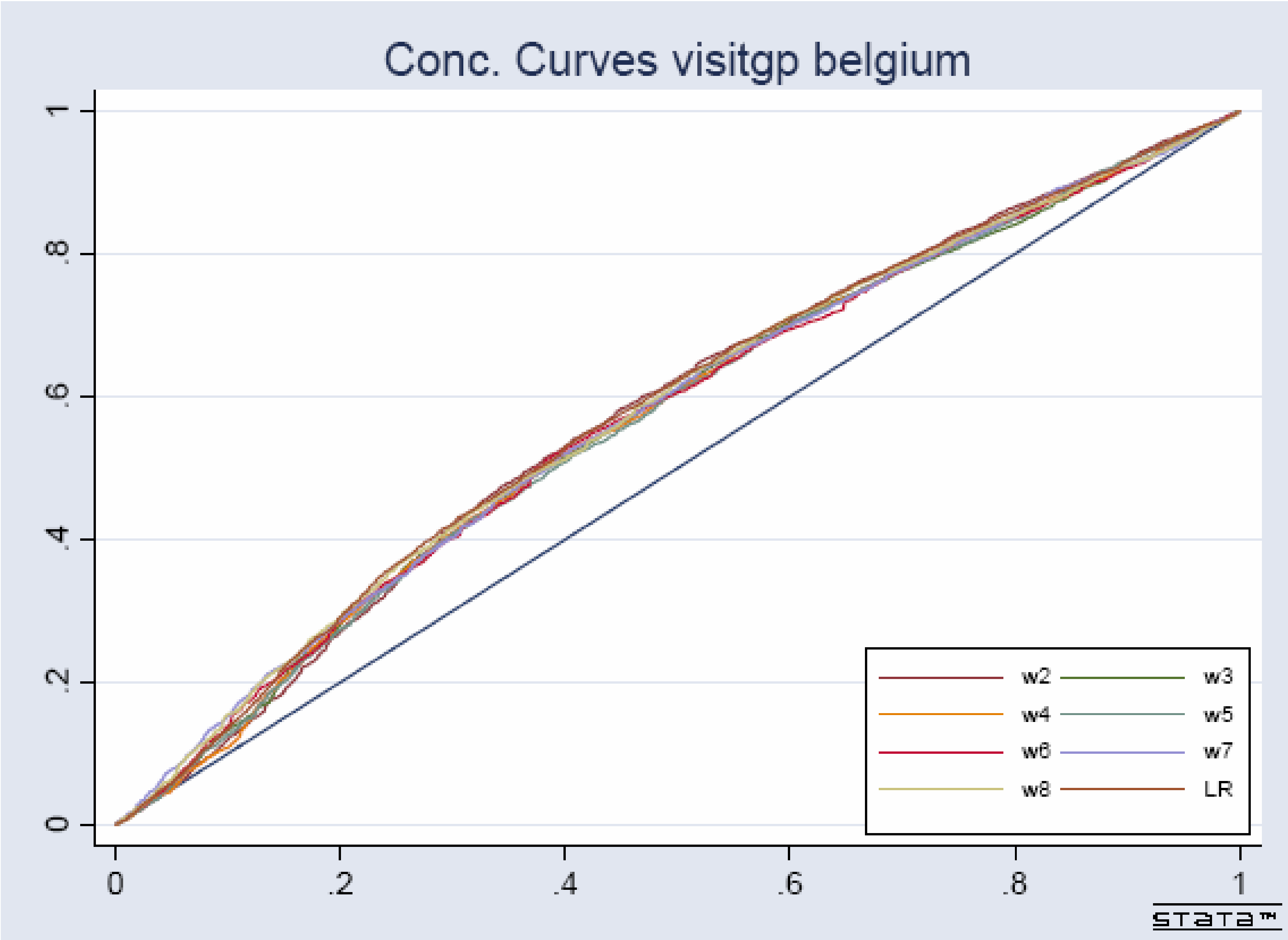

Conc. Curves visitgp belgium
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™

Figure 2

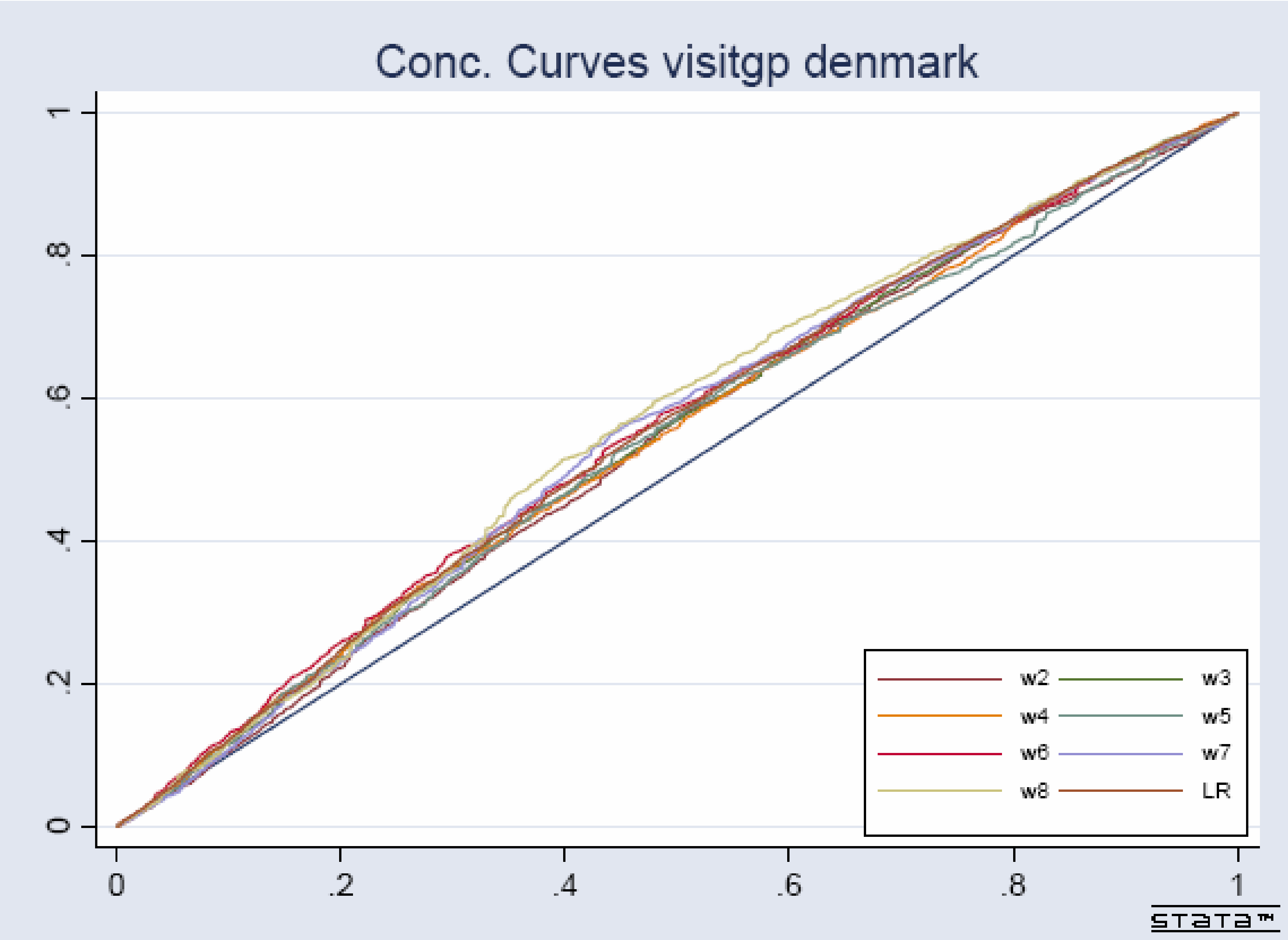

Conc. Curves visitgp denmark
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™

Figure 3

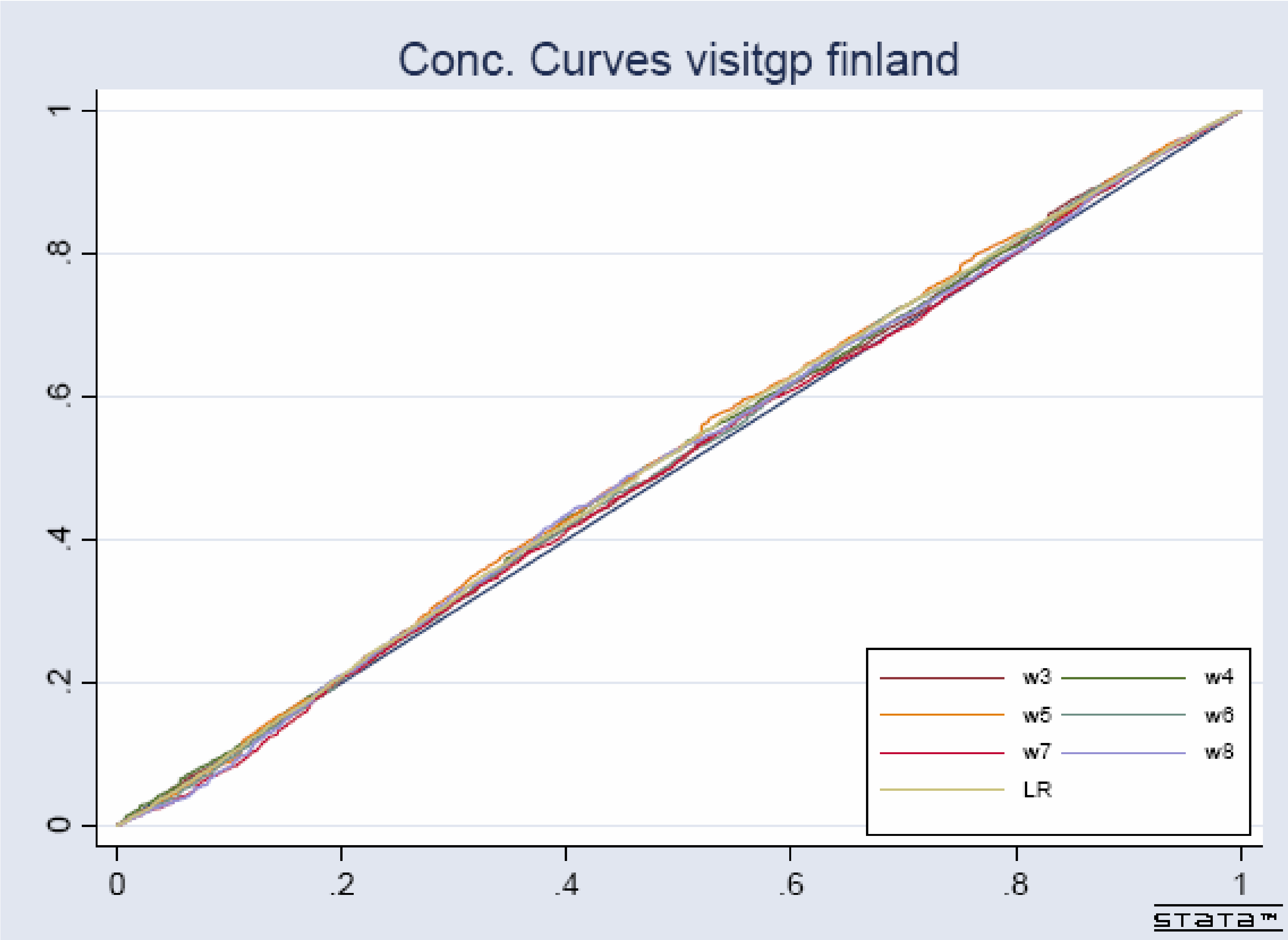


Figure 4

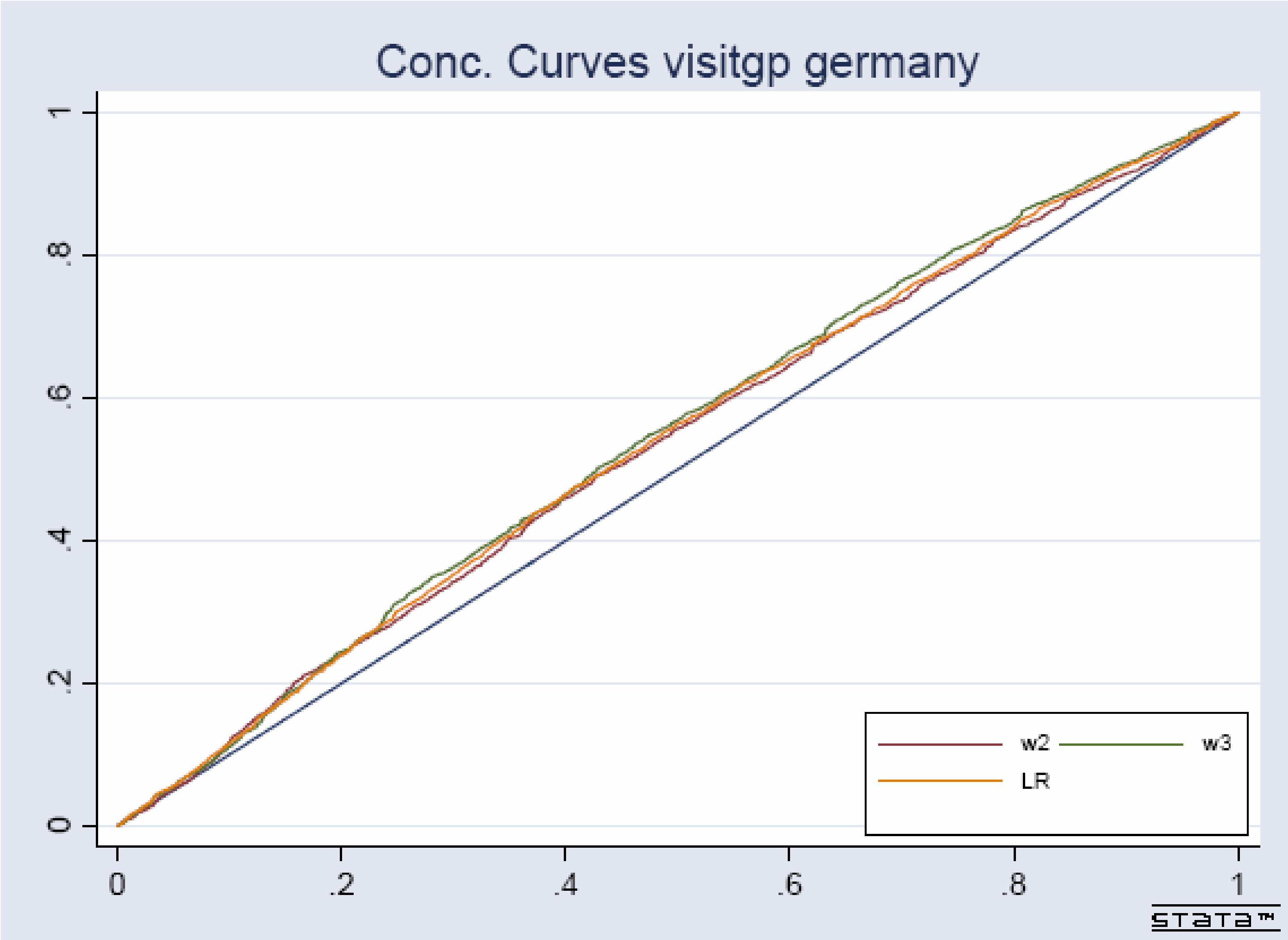
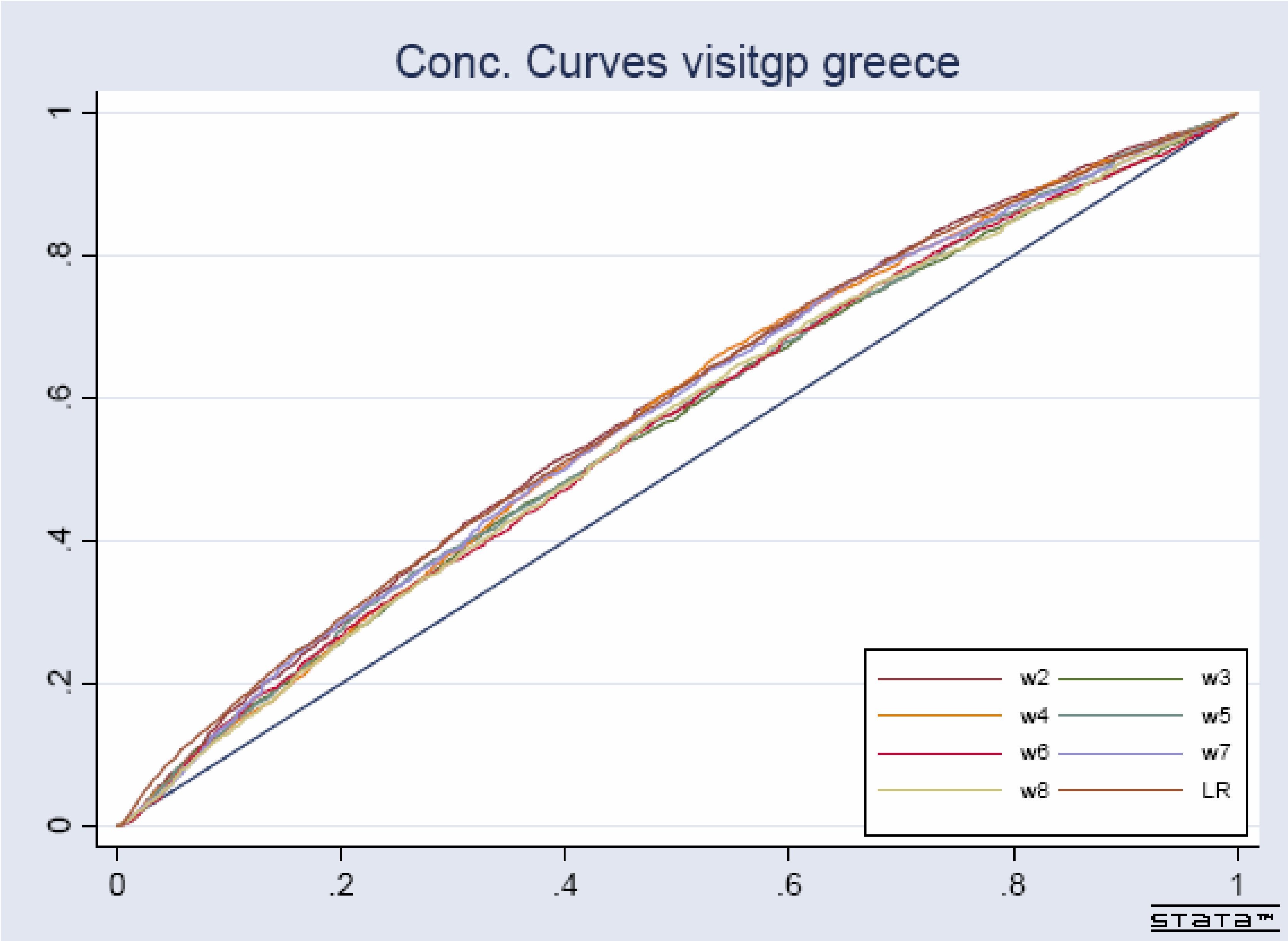


Figure 5


Conc. Curves visitgp greece
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™
Figure 6

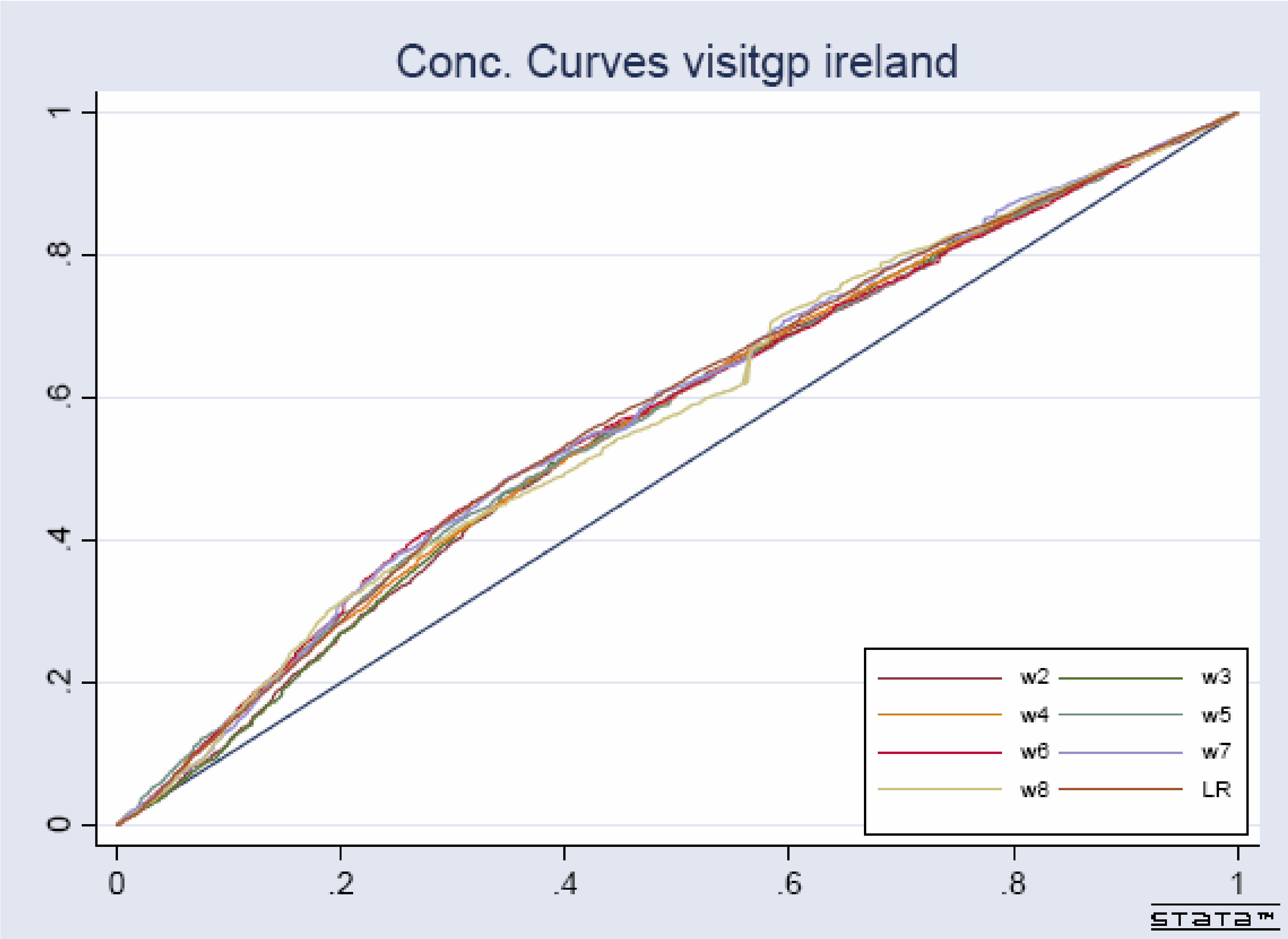

Conc. Curves visitgp ireland
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR

Figure 7

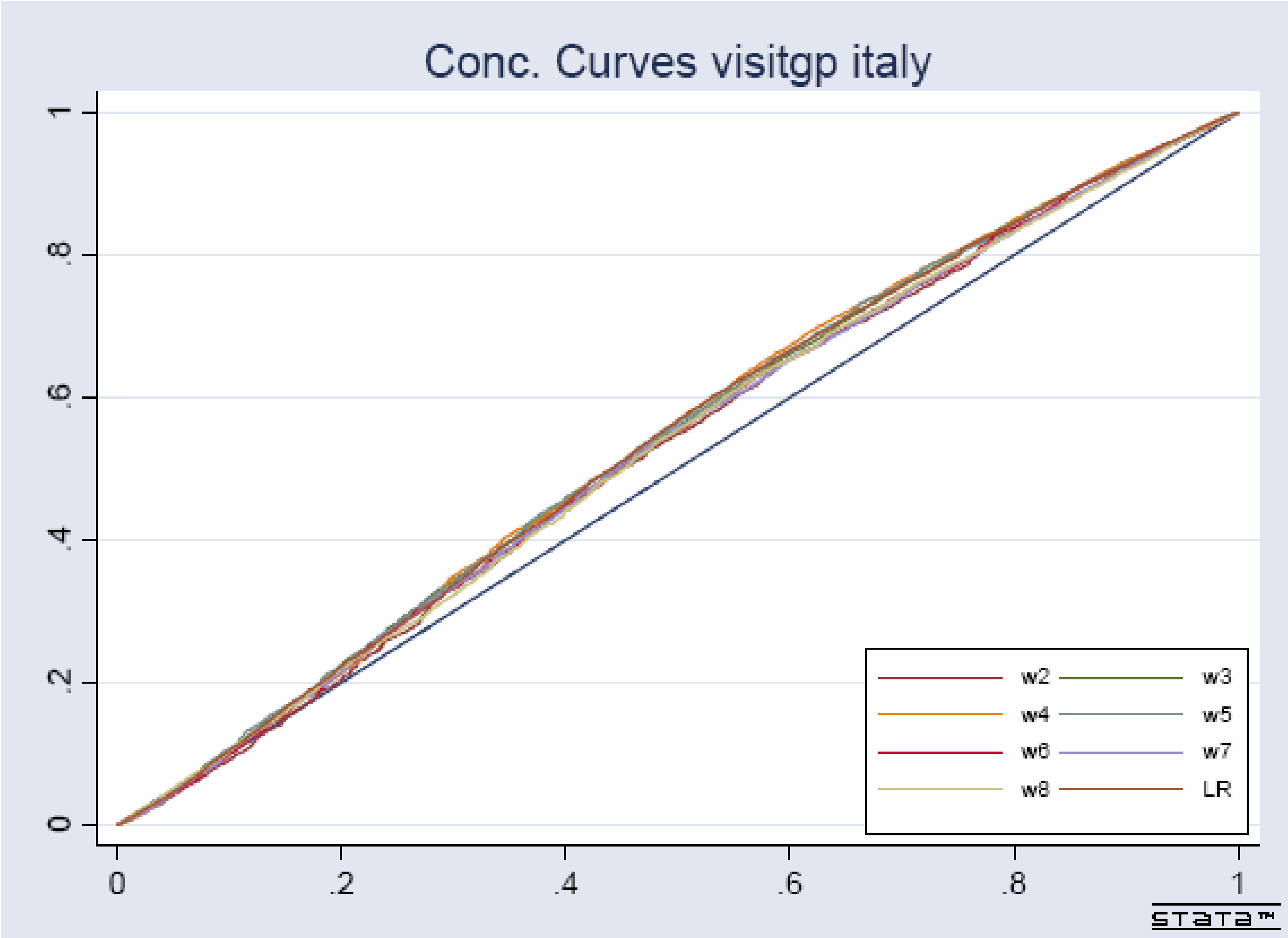

Conc. Curves visitgp italy
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™

Figure 8

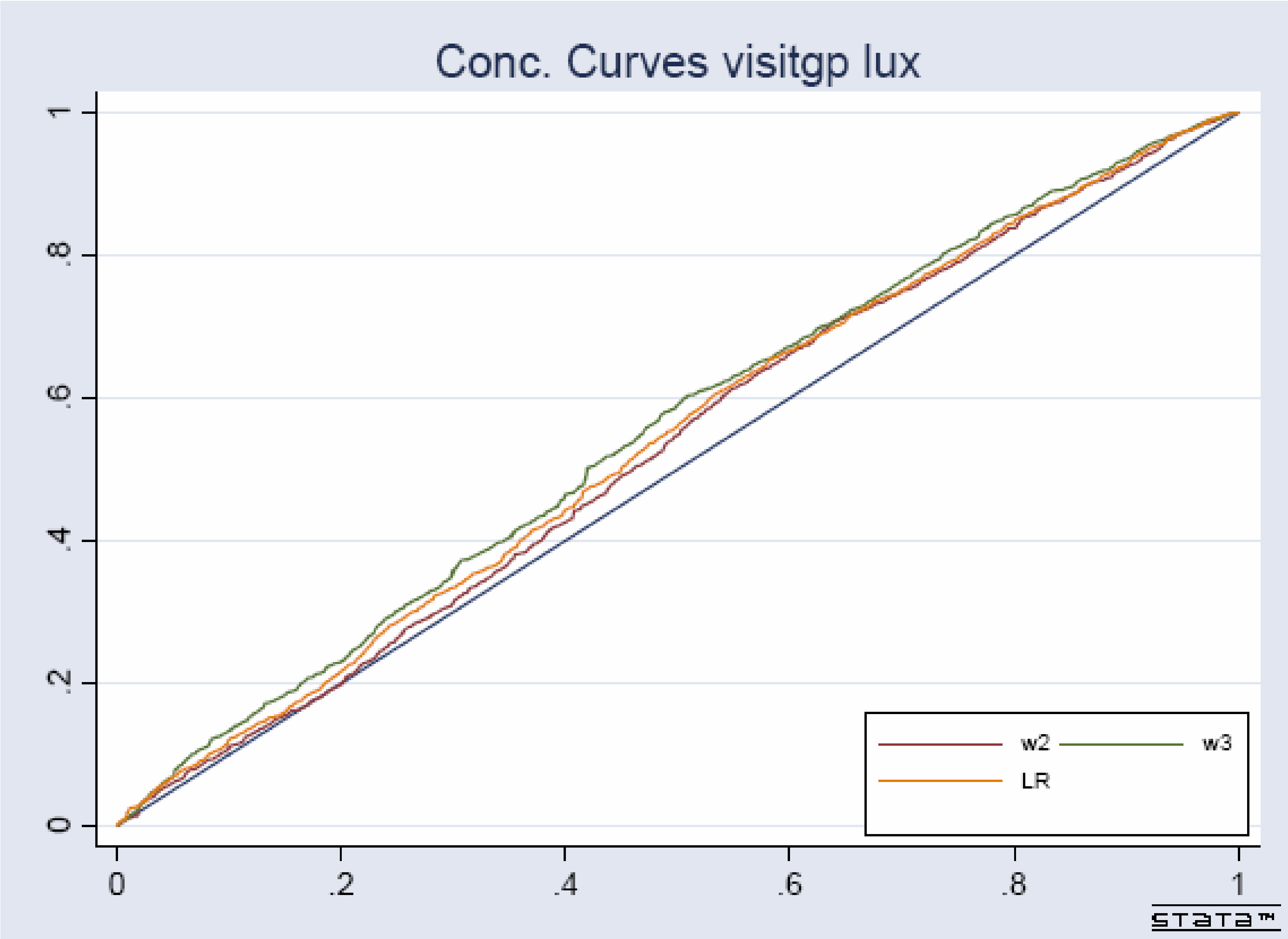
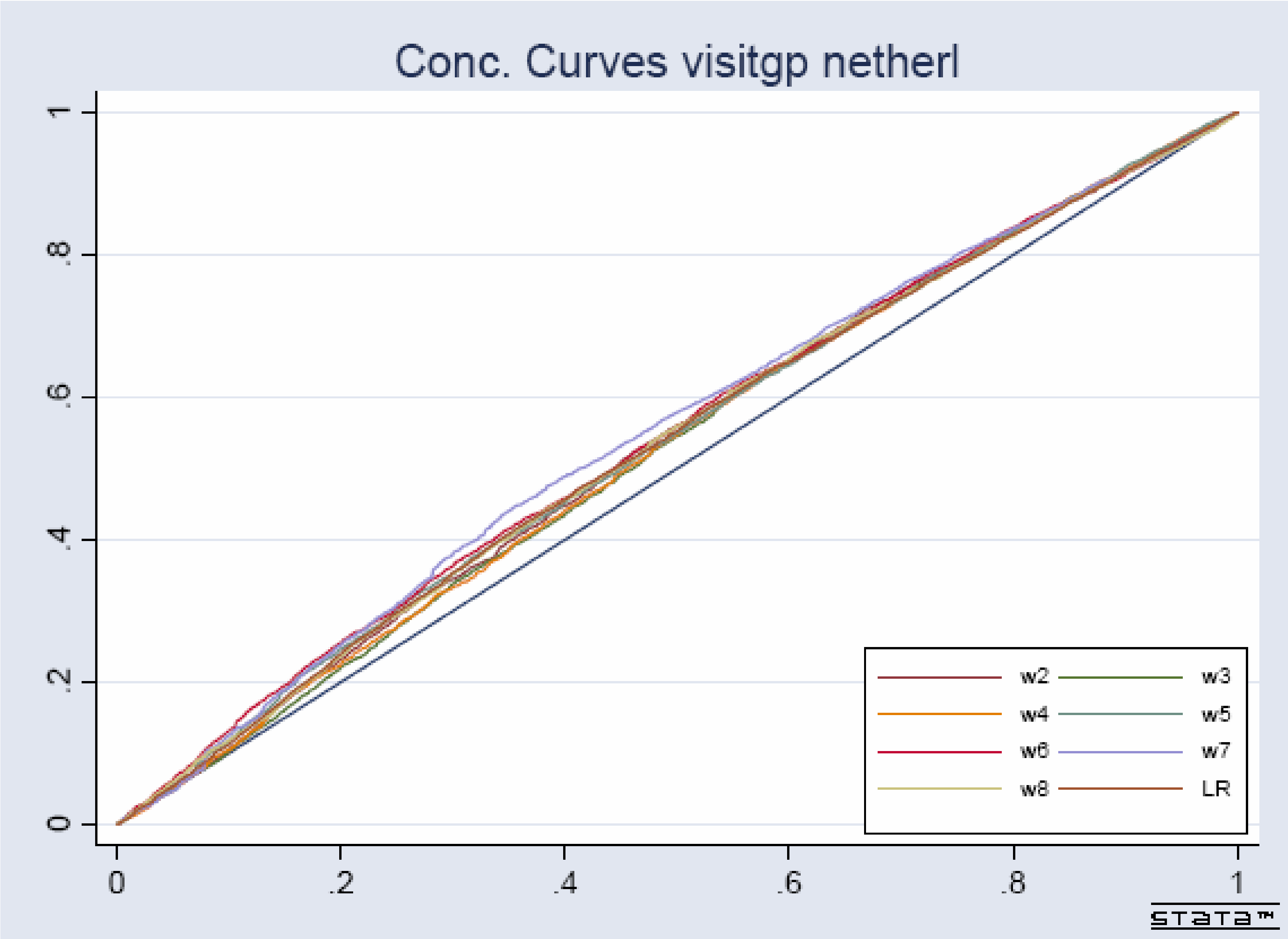
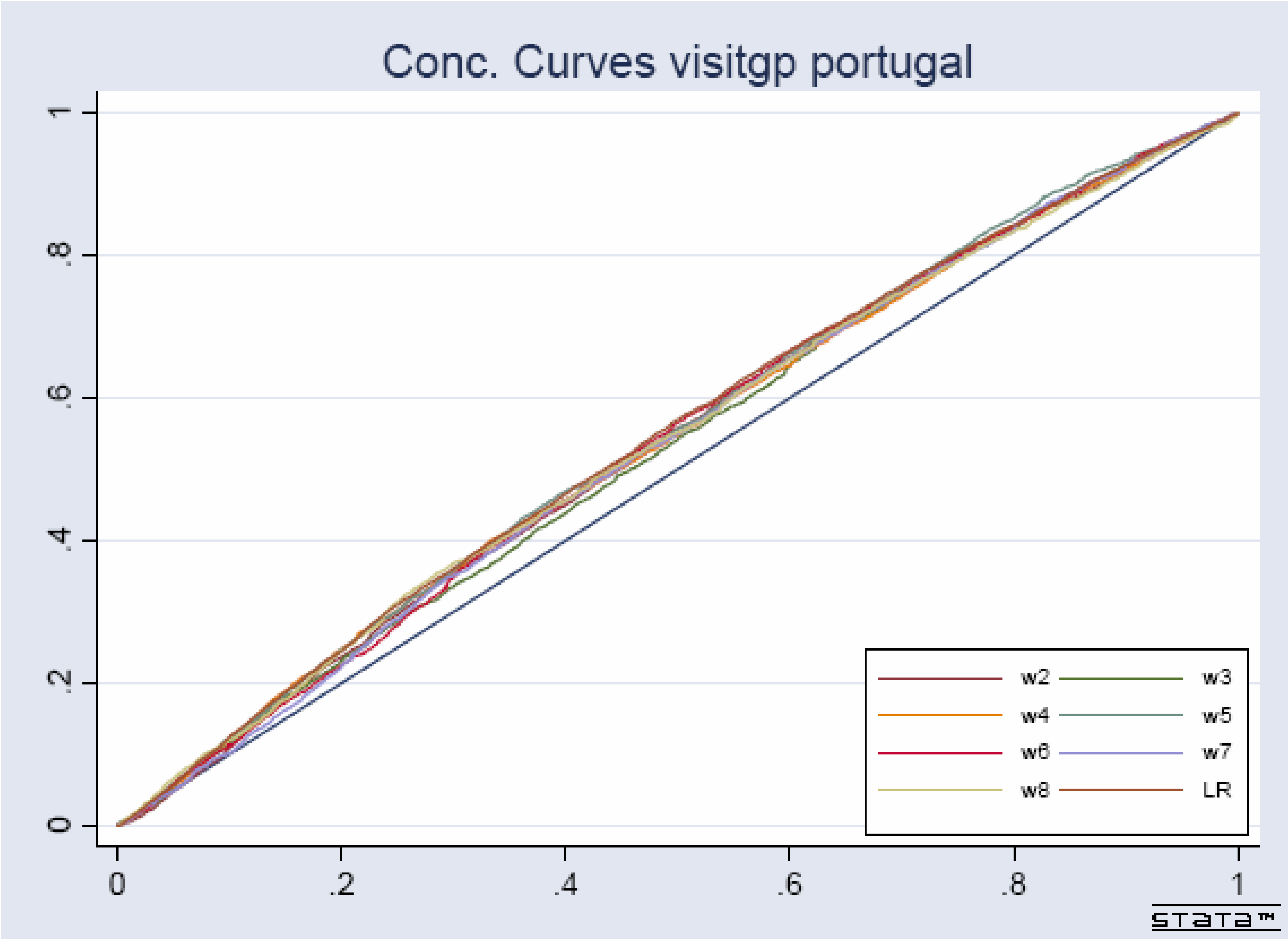


Figure 9


Conc. Curves visitgp netherl
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™
Figure 10


Conc. Curves visitgp portugal
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™
Figure 11

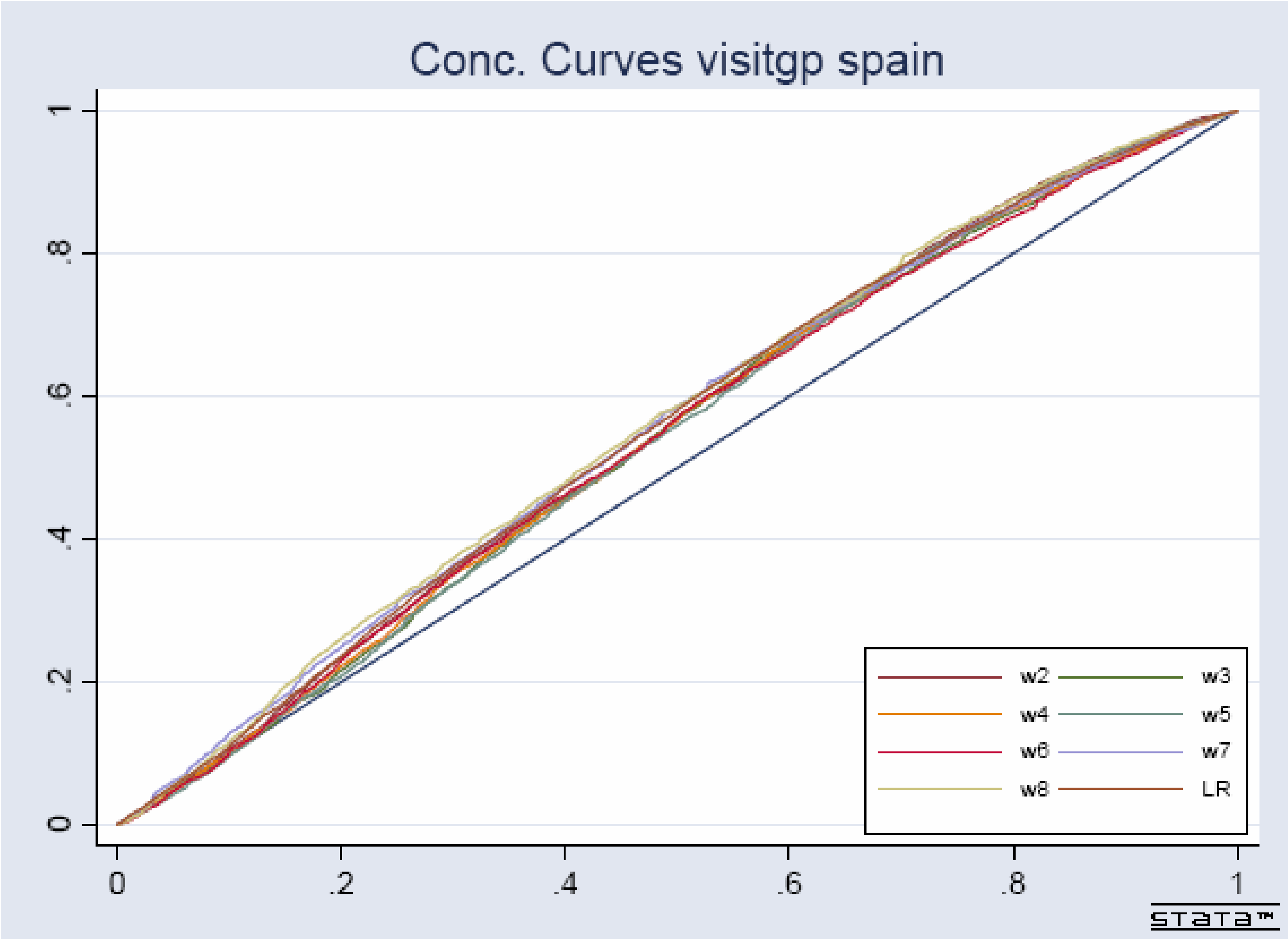

Conc. Curves visitgp spain
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™

Figure 12

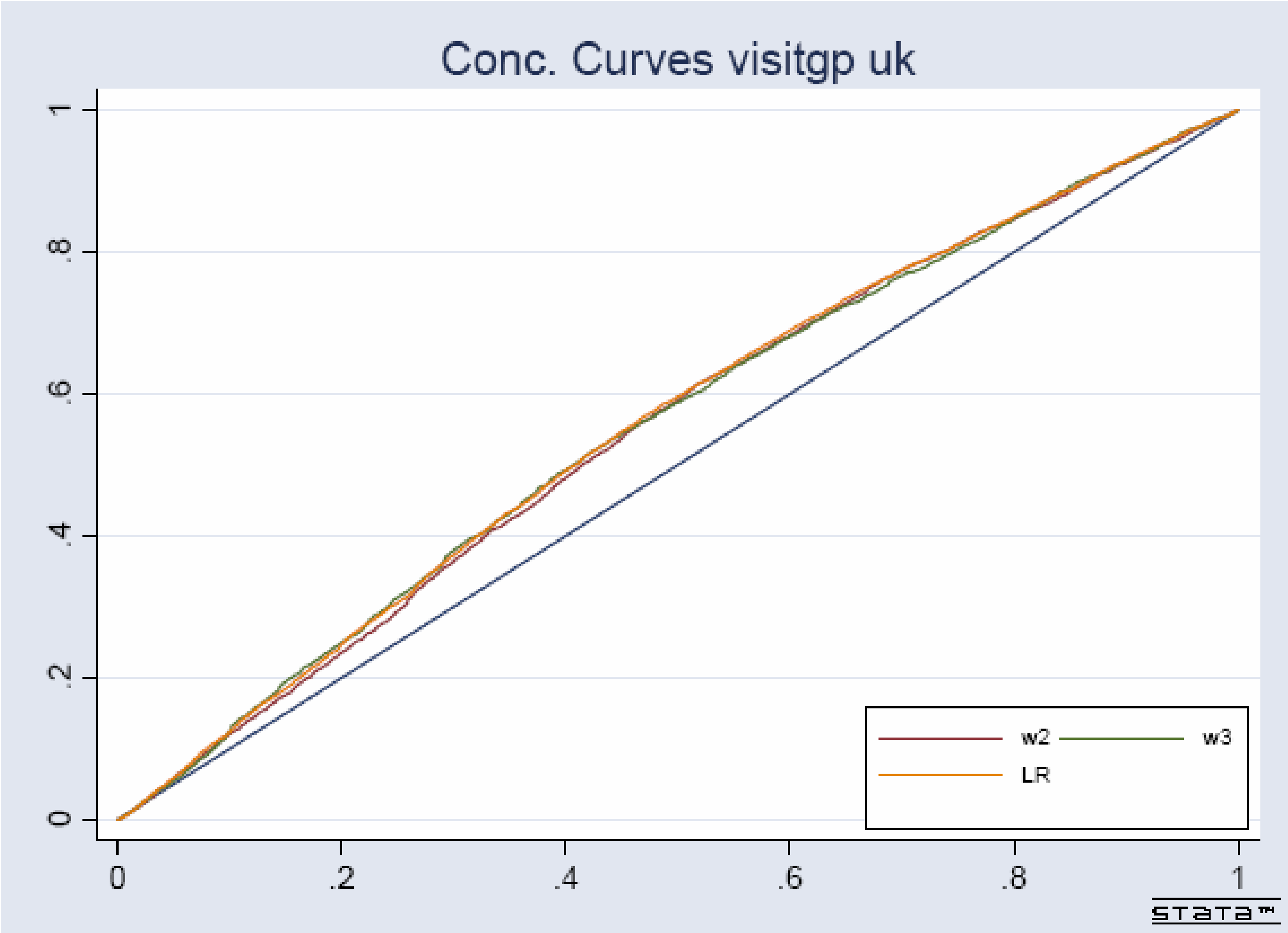
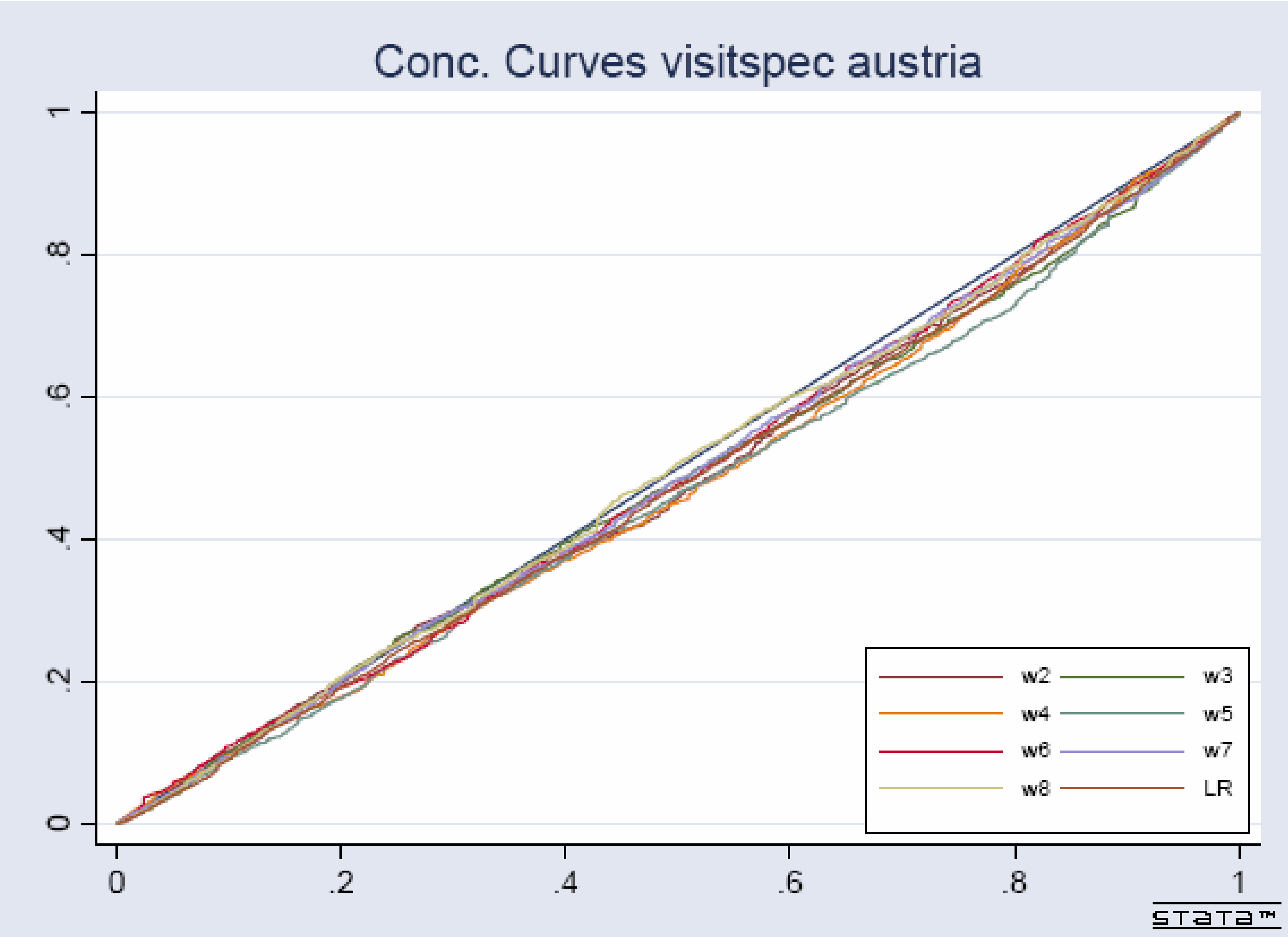
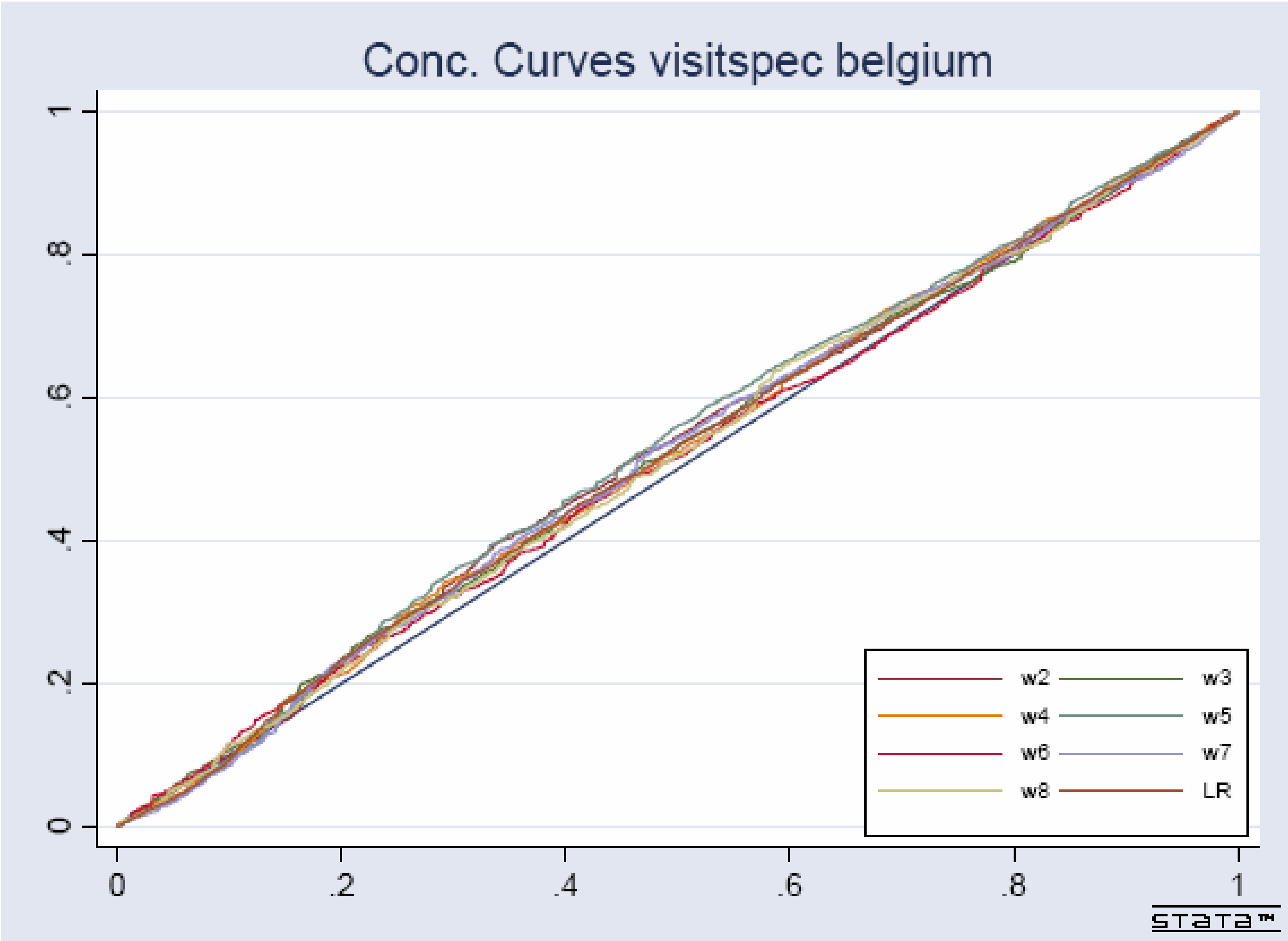
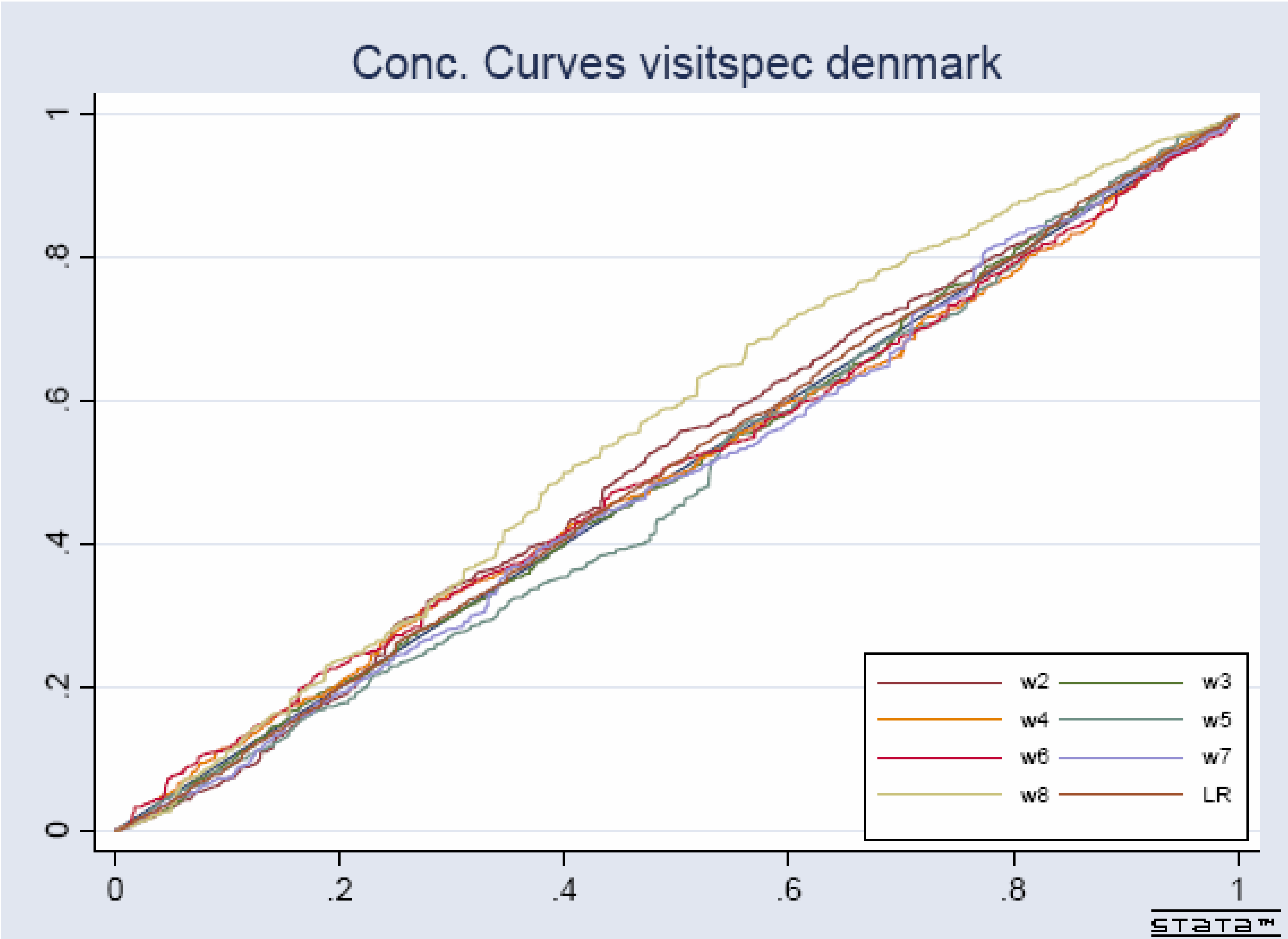
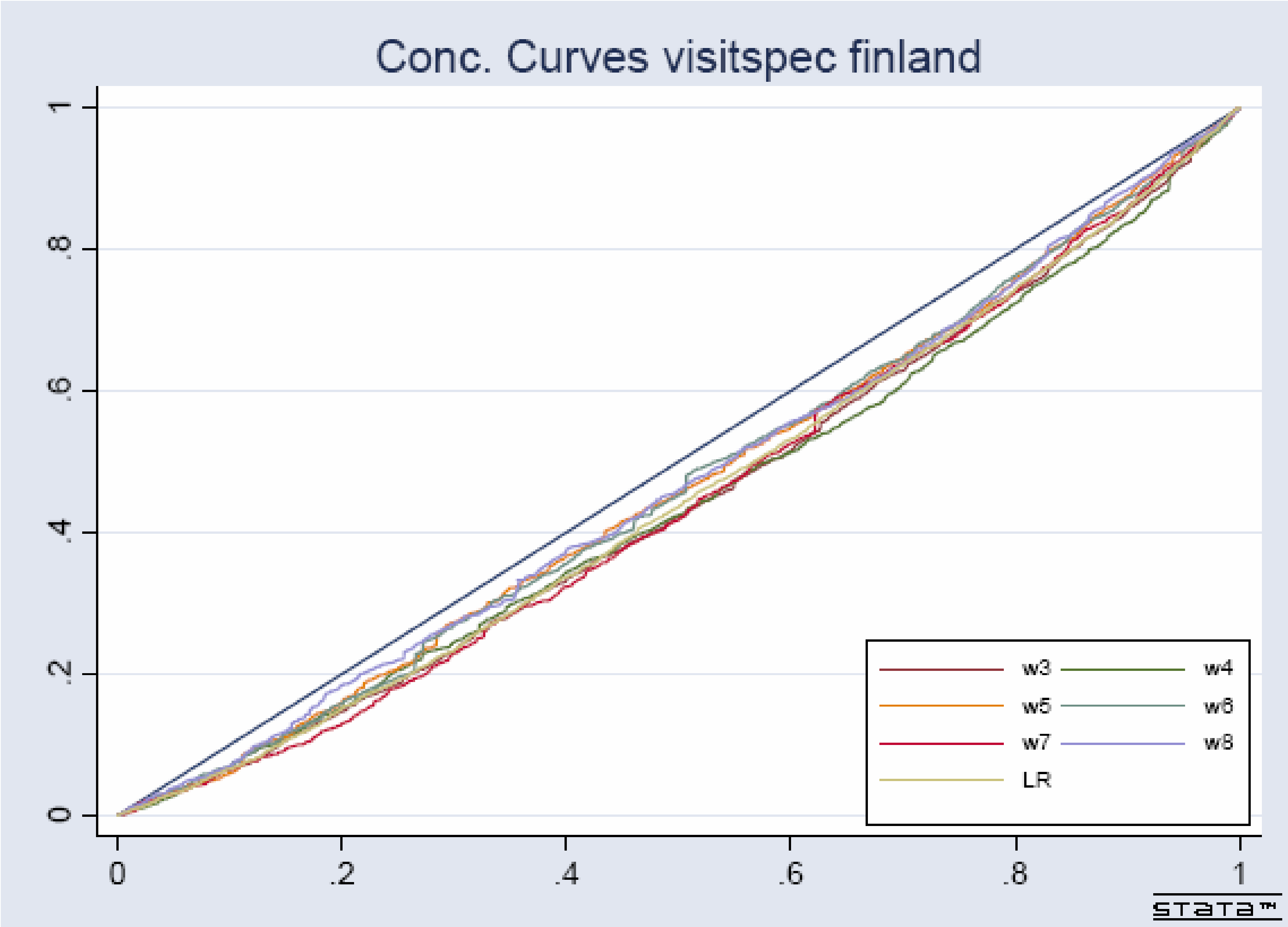


Figure 13


Conc. Curves visitspec austria
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™
Figure 14


Conc. Curves visitspec belgium
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™
Figure 15


Conc. Curves visitspec denmark
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™
Figure 16


Conc. Curves visitspec finland
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w3
w4
w5
w6
w7
w8
LR
stata™
Figure 17

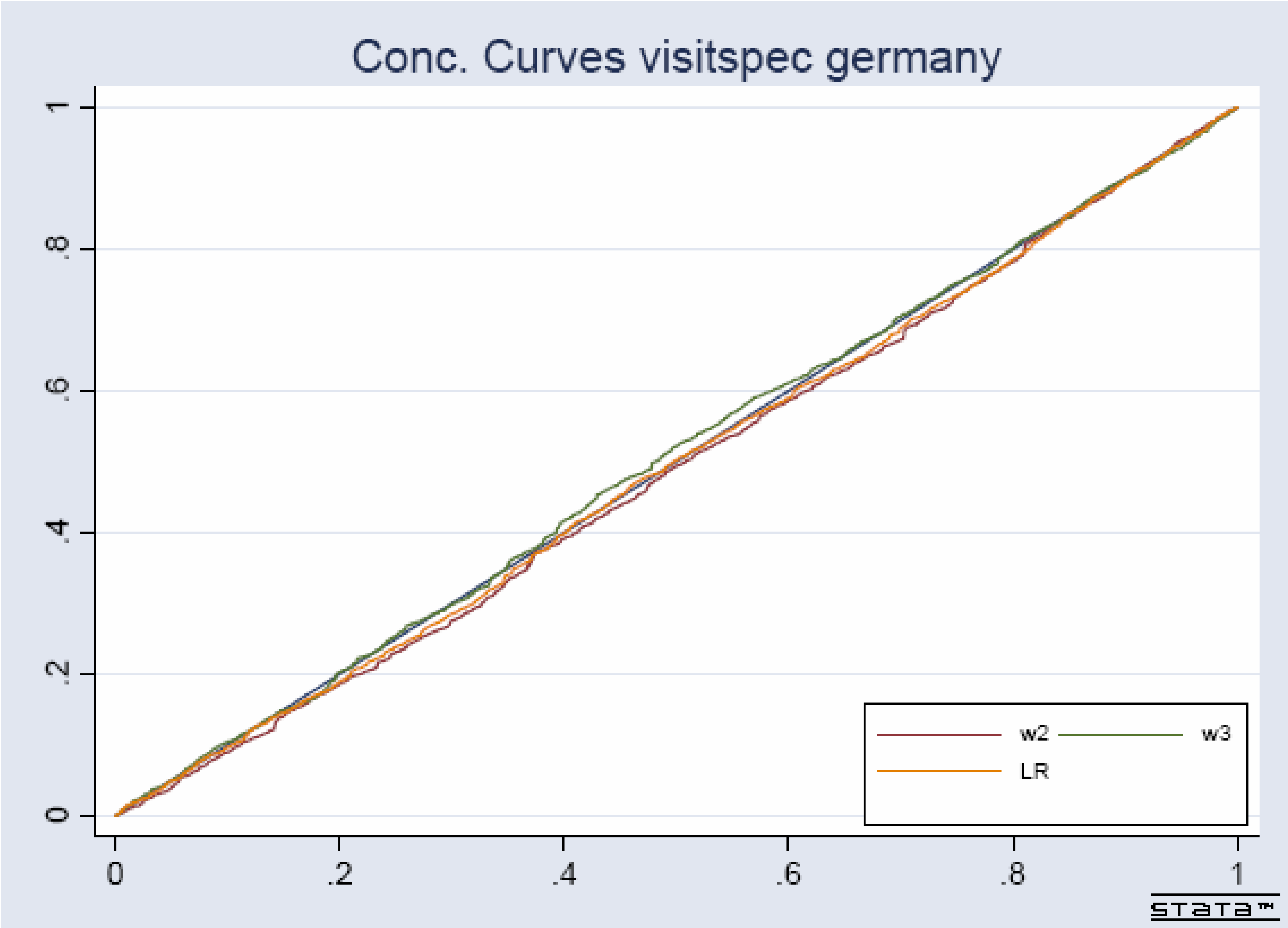
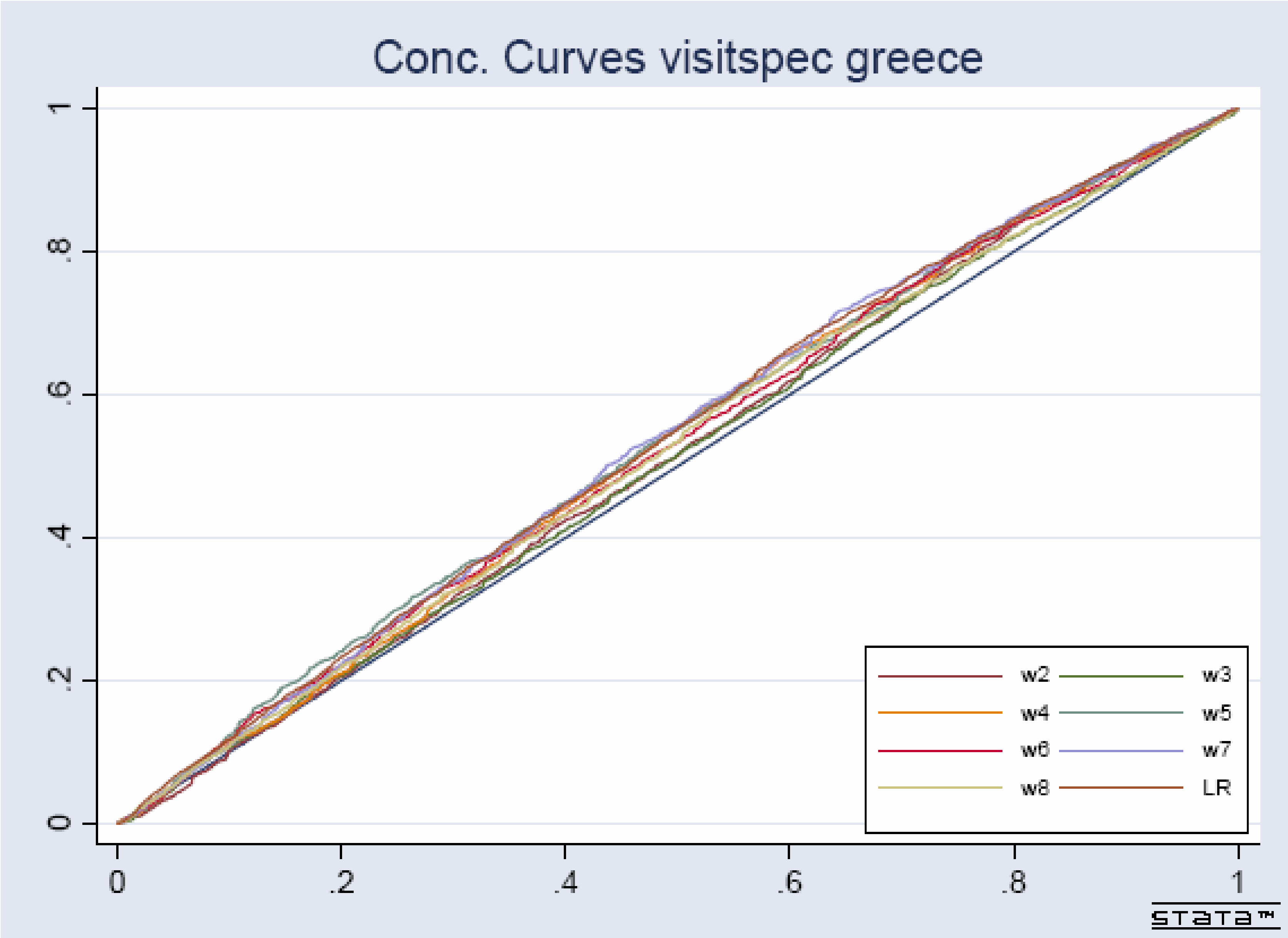


Figure 18



Conc. Curves visitspec greece
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™

Figure 19

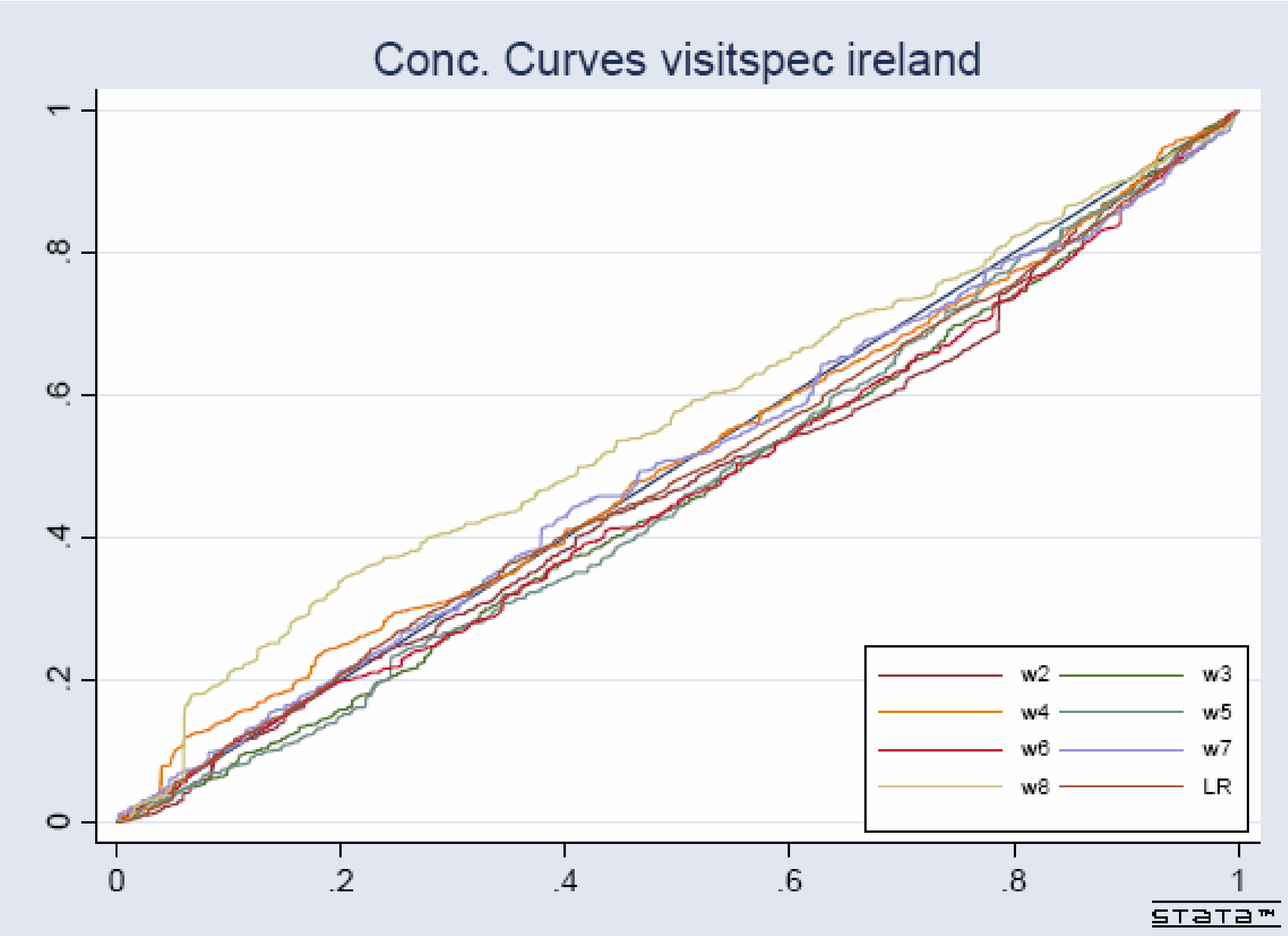
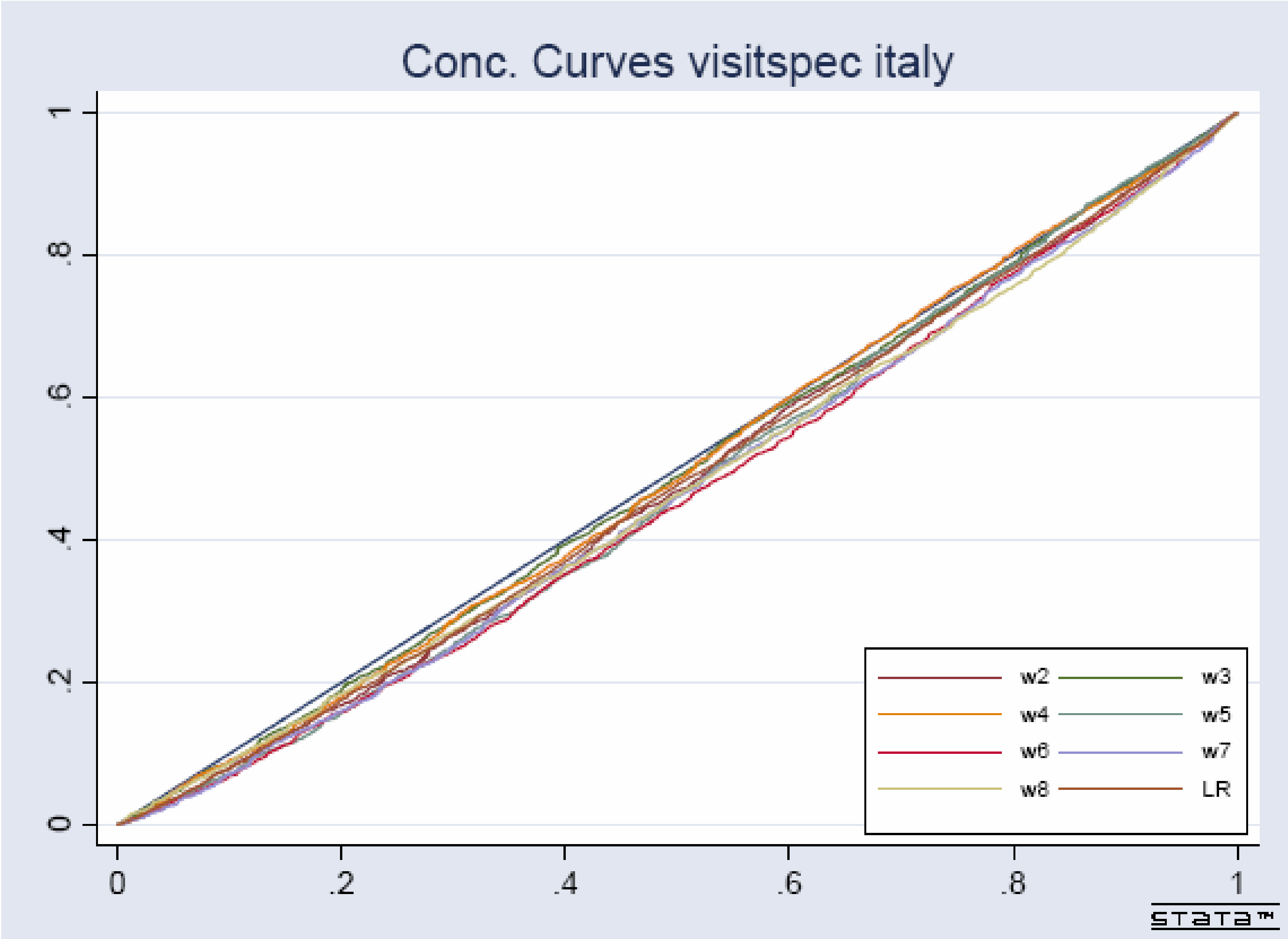

Conc. Curves visitspec ireland
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR

Figure 20


Conc. Curves visitspec italy
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™
Figure 21

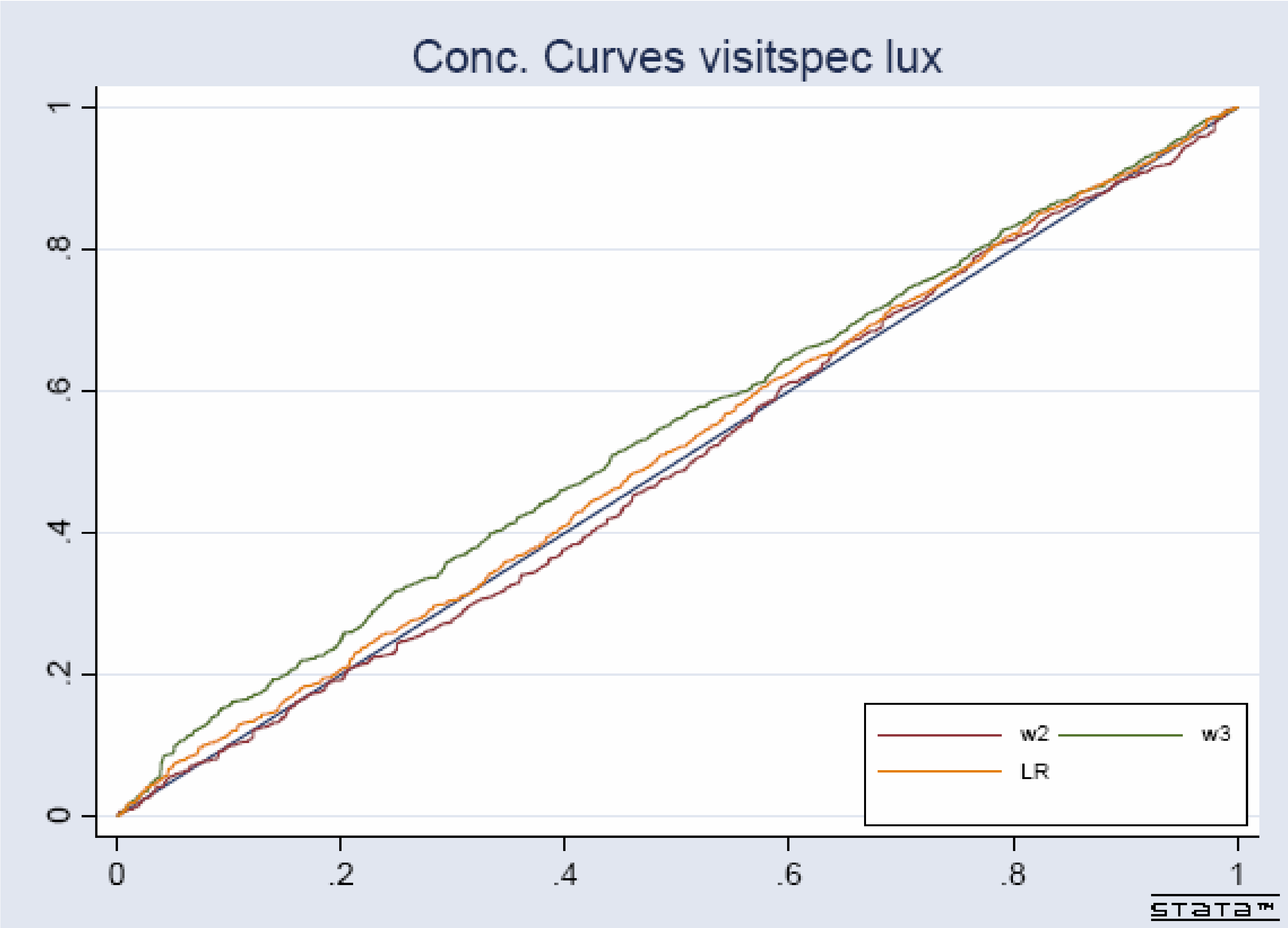
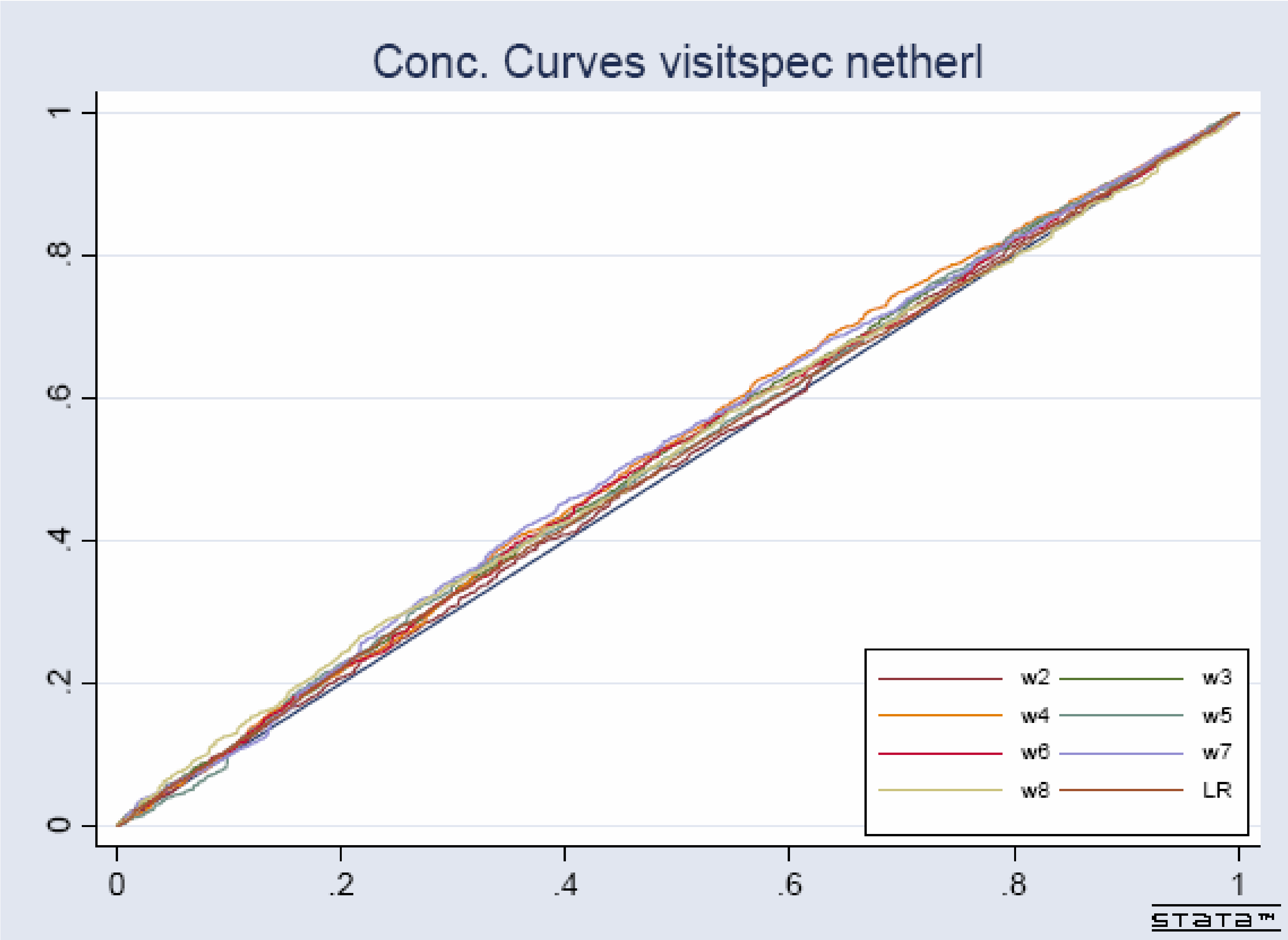


Figure 22


Conc. Curves visitspec netherl
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™
Figure 23

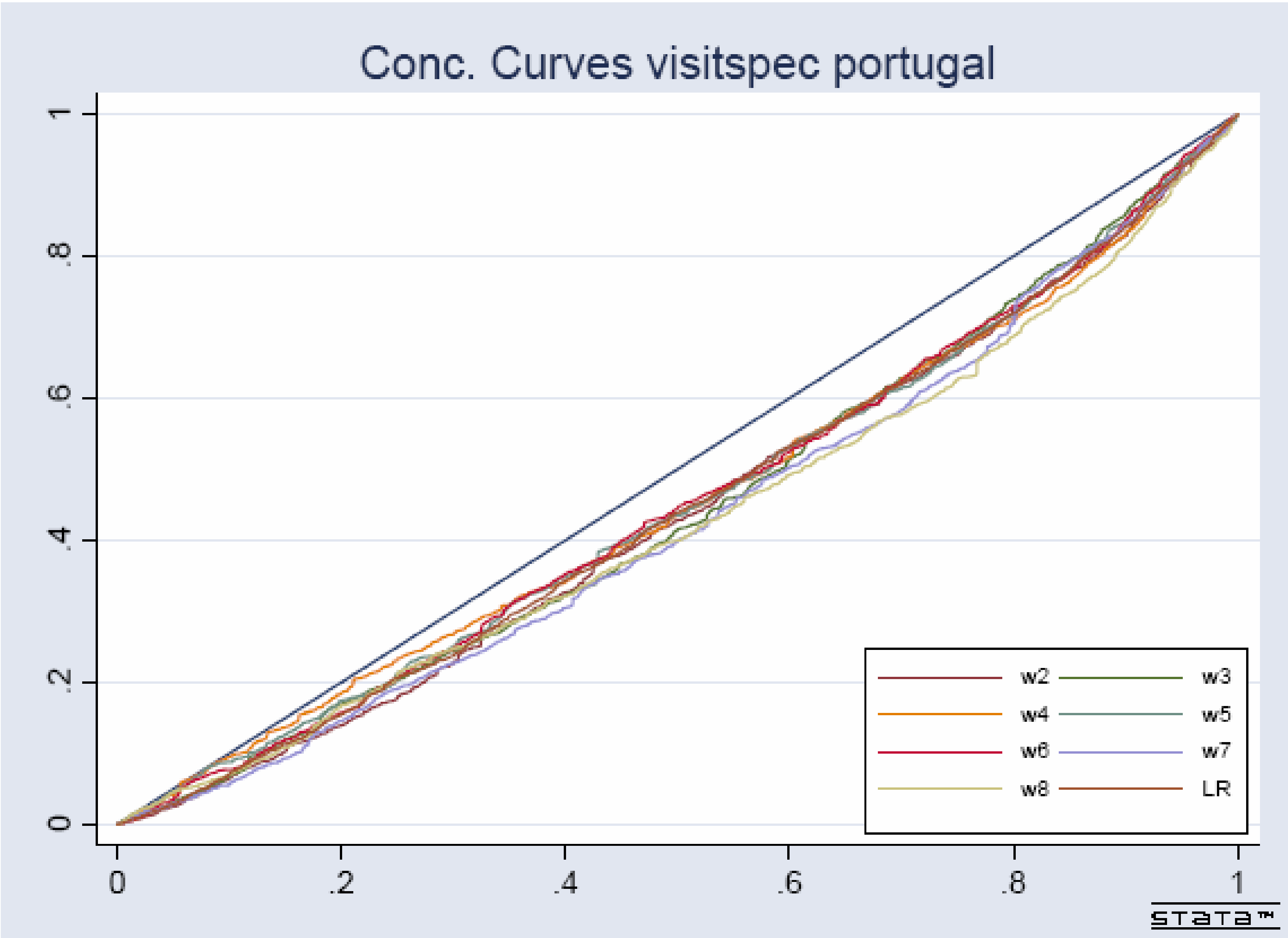

Conc. Curves visitspec portugal
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™

Figure 24

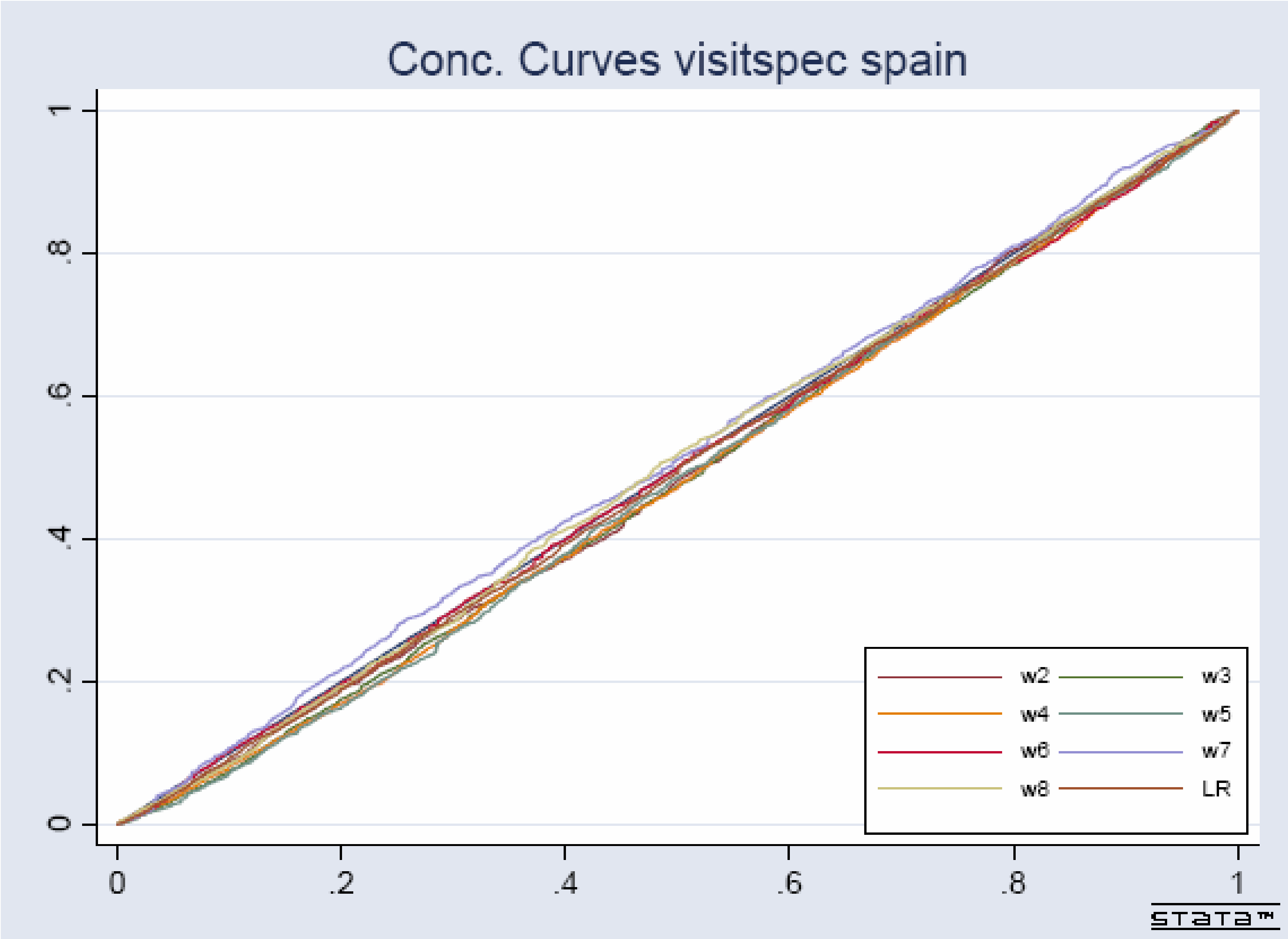

Conc. Curves visitspec spain
0
.2
.4
.6
.8
1
0
.2
.4
.6
.8
1
w2
w3
w4
w5
w6
w7
w8
LR
stata™

Figure 25

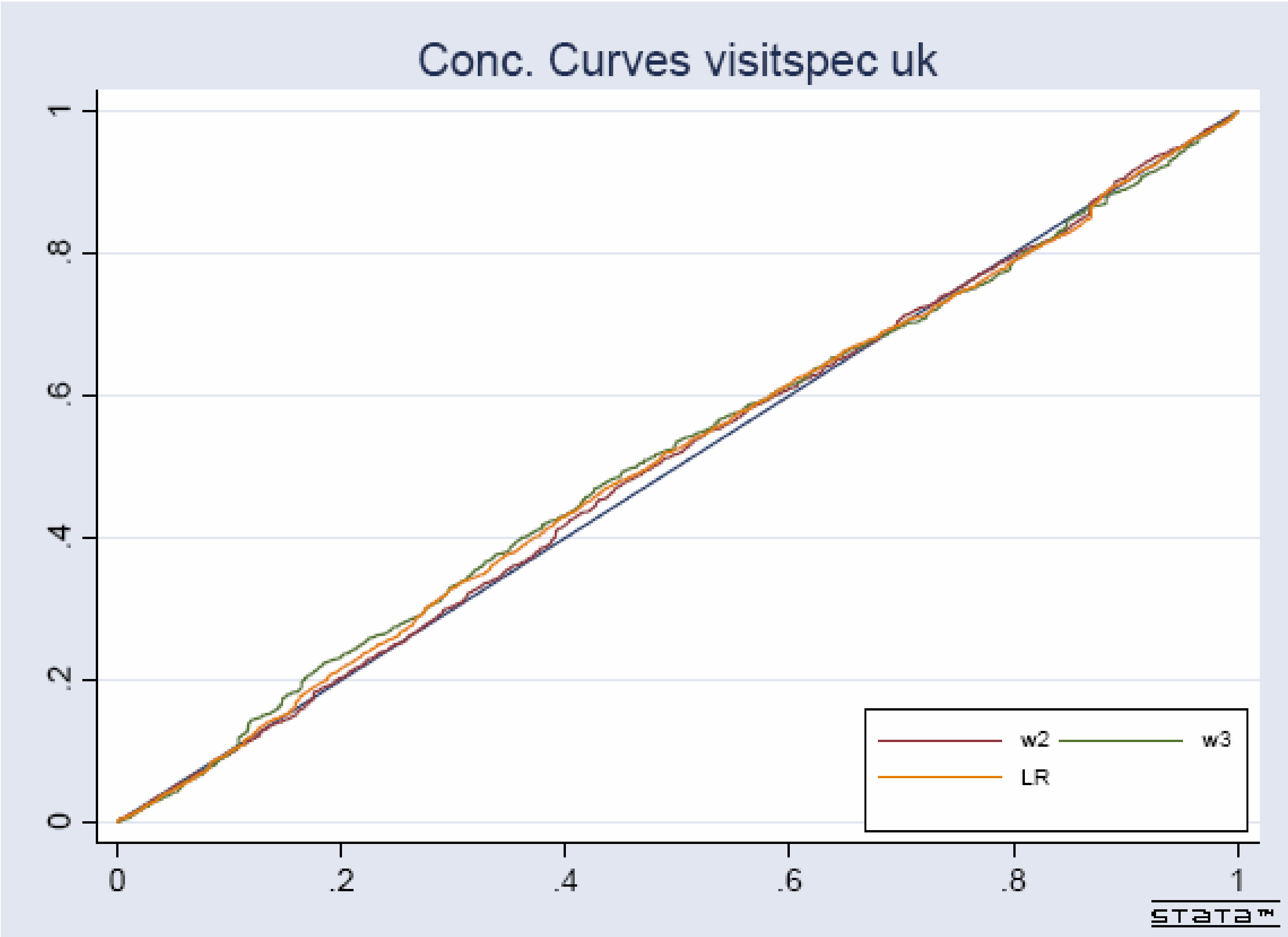


Figure 26

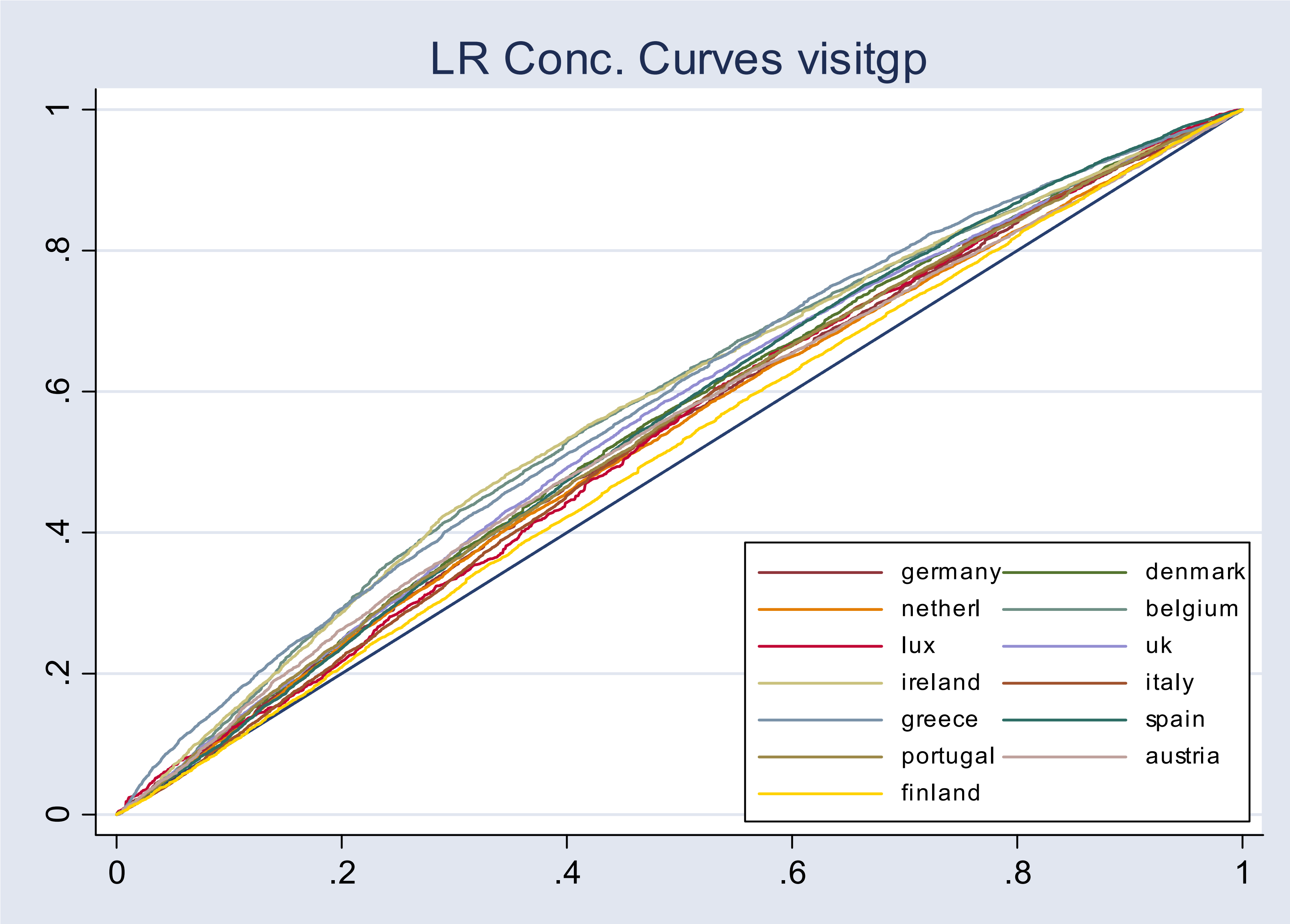


Figure 27: Long-run concentration curves of GP visits (all countries)

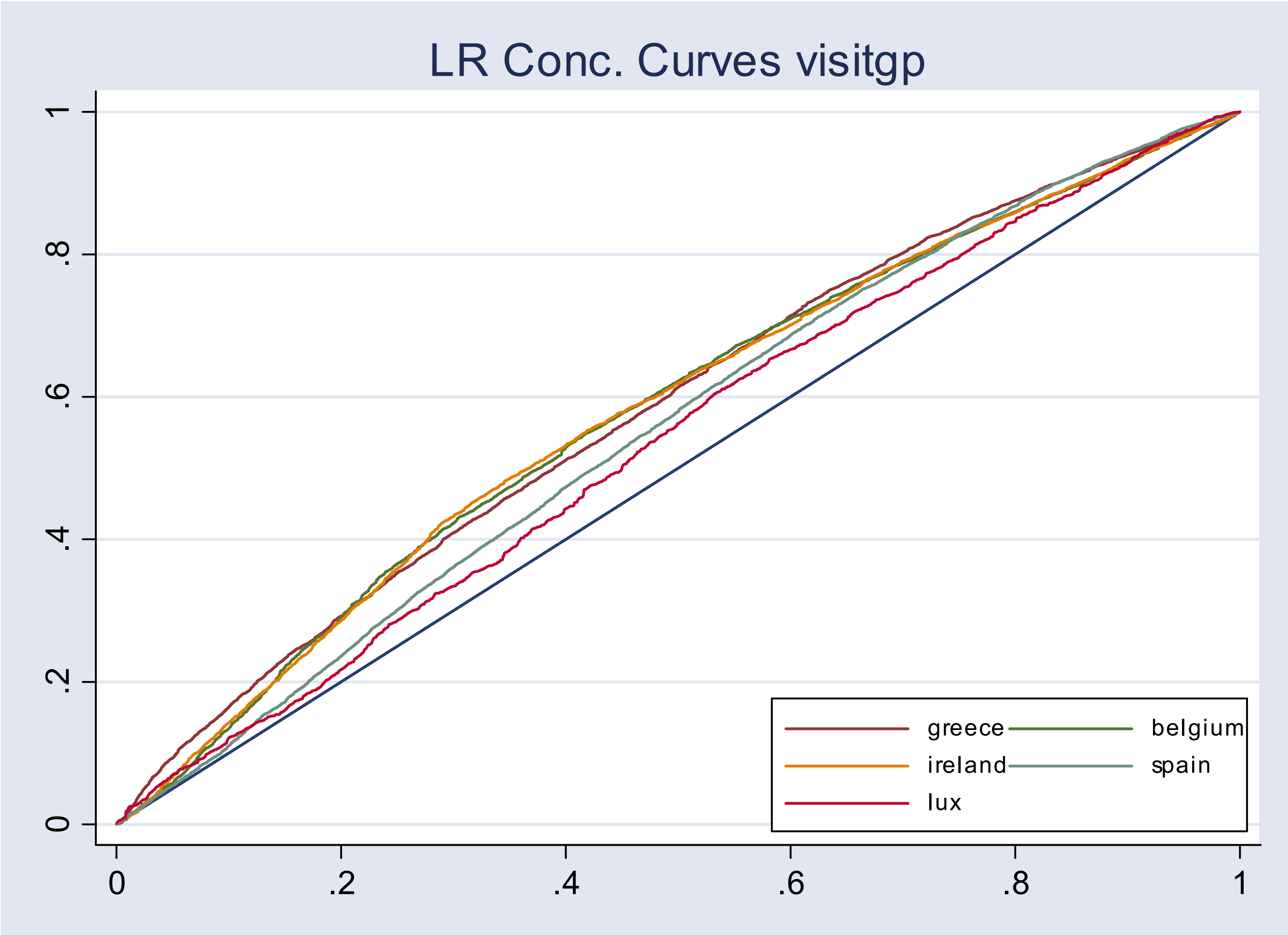


Figure 28: Long-run concentration curves of GP visits (1st selection)

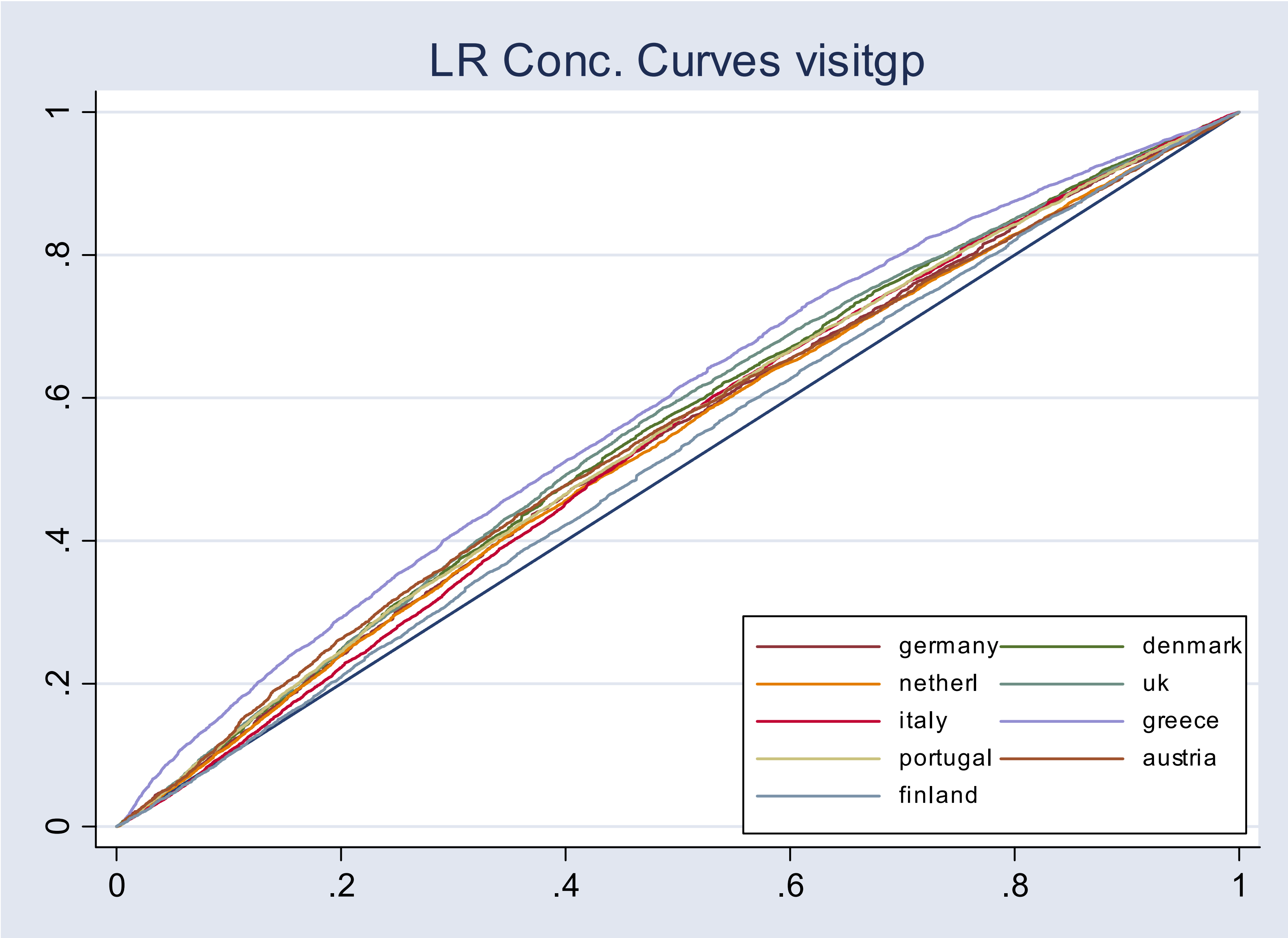


Figure 29: Long-run concentration curves of GP visits (2nd selection)

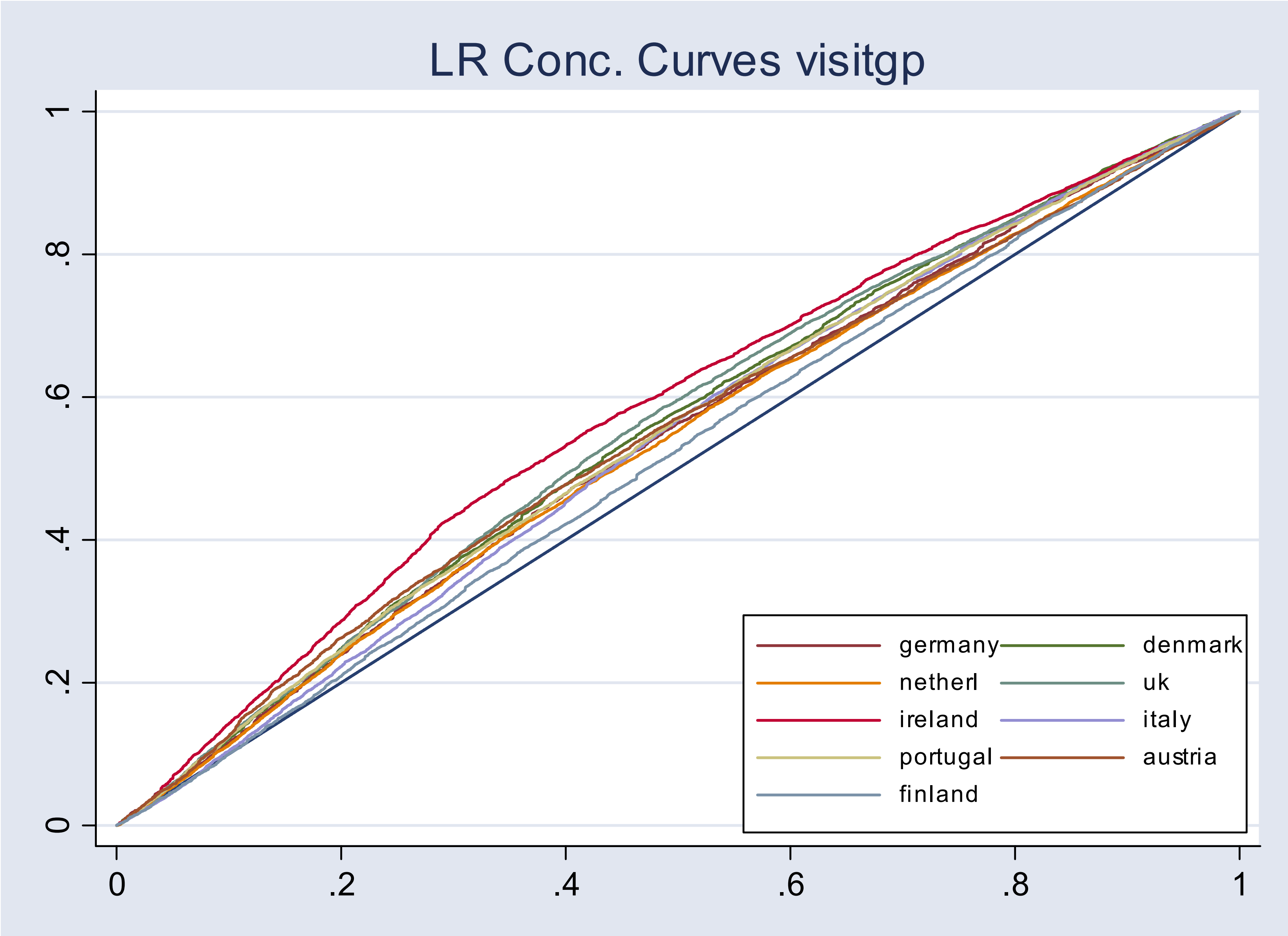


Figure 30: Long-run concentration curves of GP visits (3rd selection)

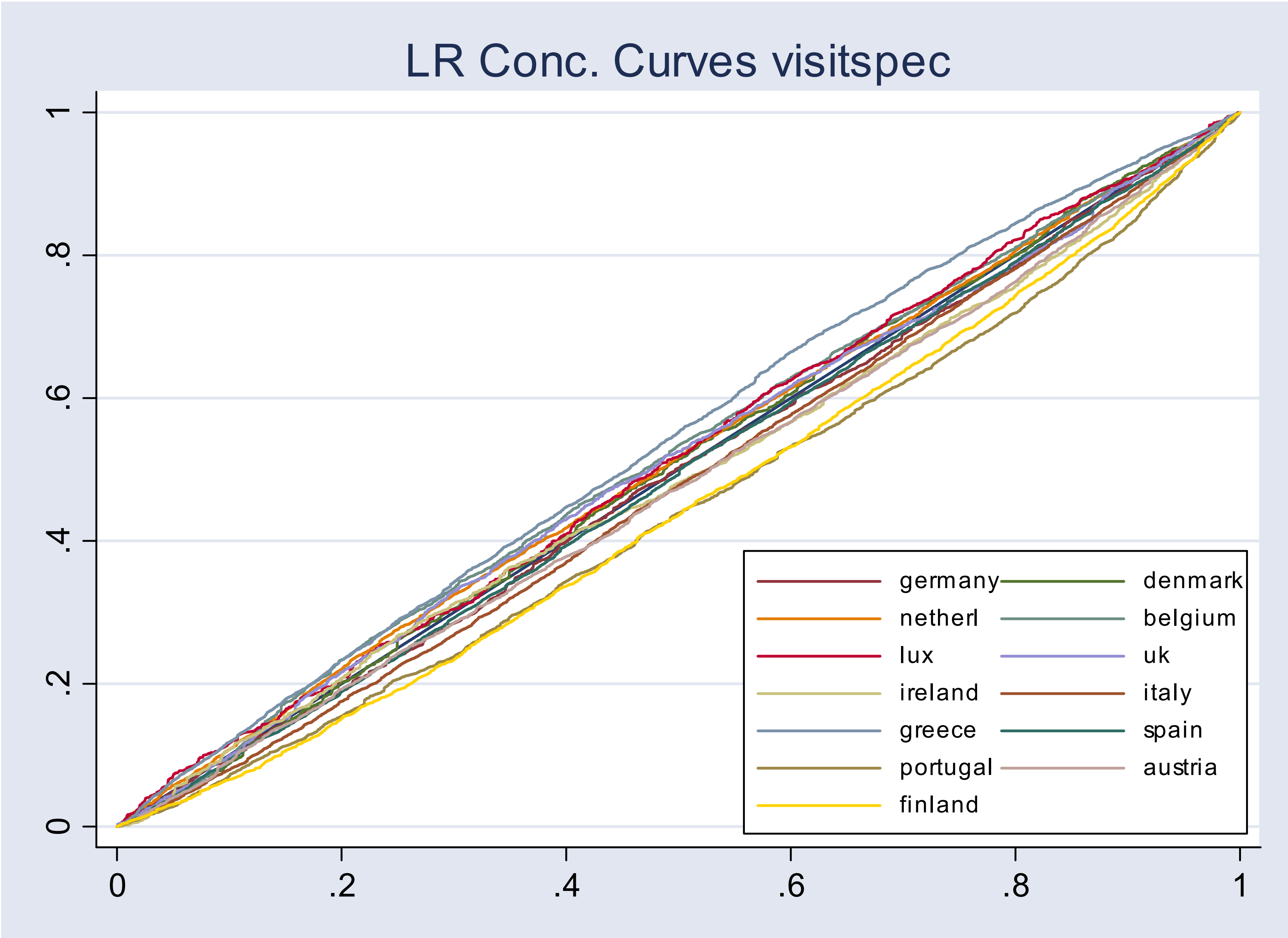


Figure 31: Long-run concentration curves of specialist visits (all countries)

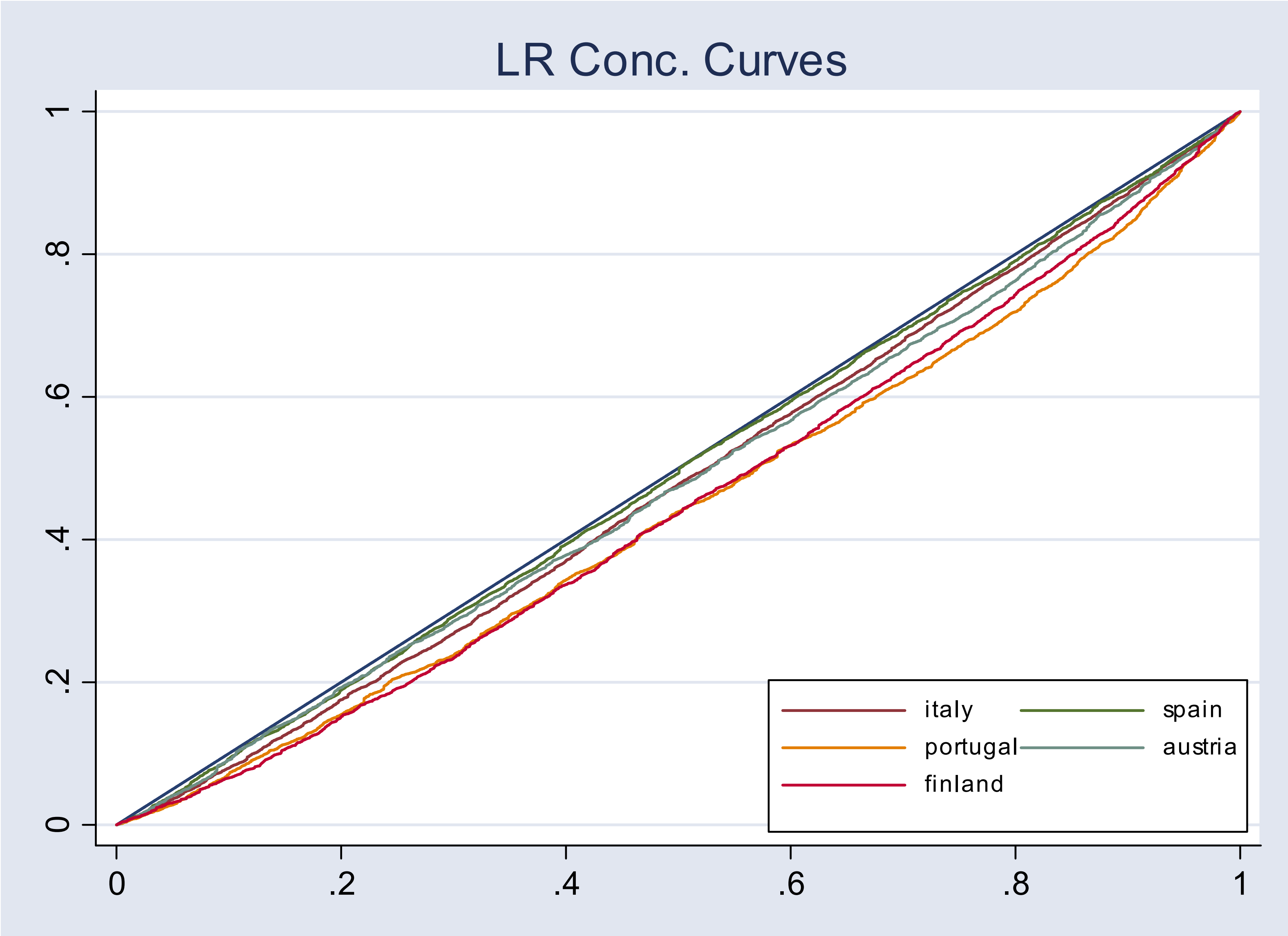


Figure 32: Long-run concentration curves of specialist visits (1st selection)

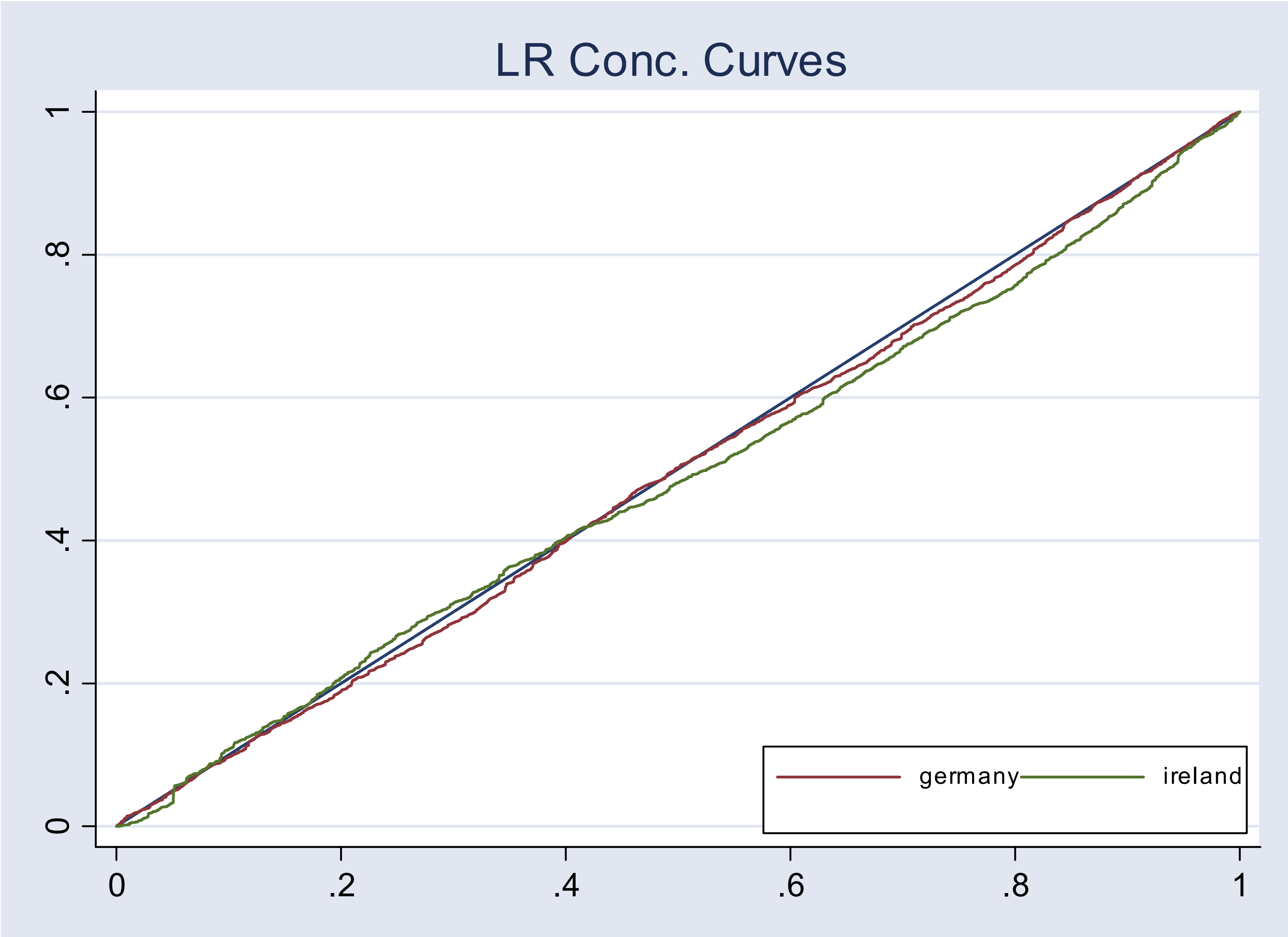


Figure 33: Long-run concentration curves of specialist visits (2nd selection)

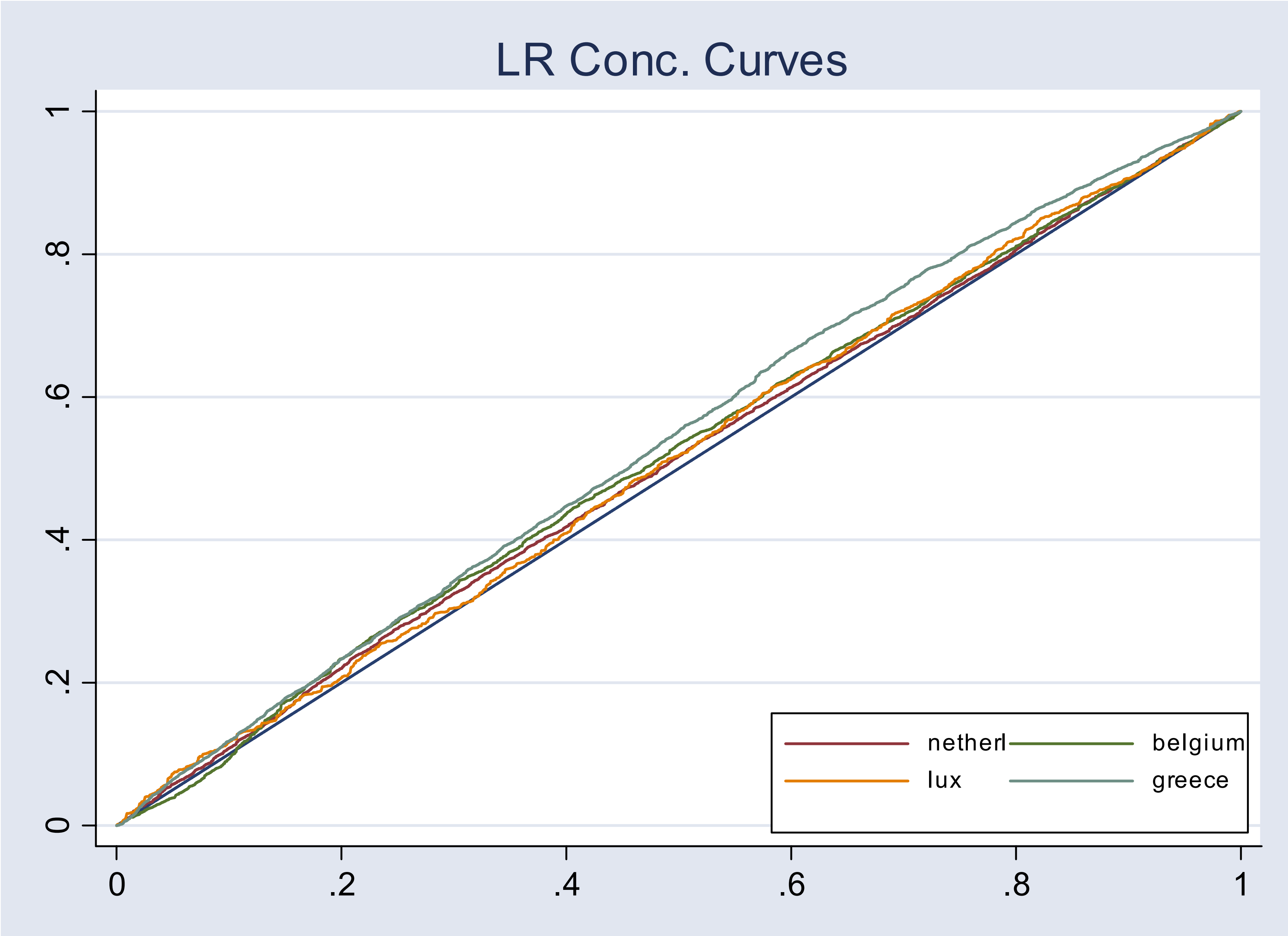


Figure 34: Long-run concentration curves of specialist visits (3rd selection)

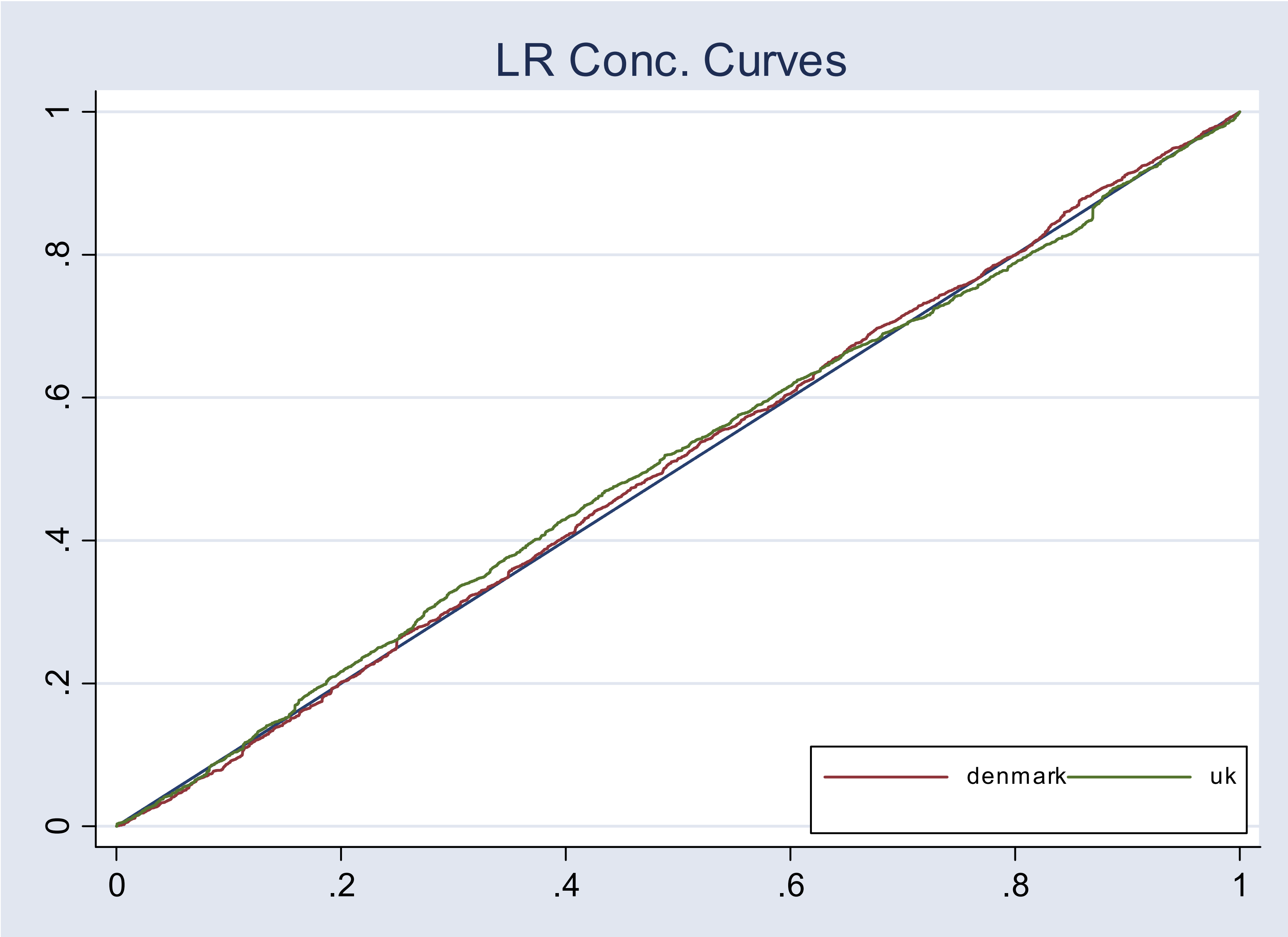


Figure 35: Long-run concentration curves of specialist visits (4th selection)