Construction of the Gender Index: Methodology

An index is a quantitative estimation of a social phenomenon, based on multiple indicators that pertain to the phenomenon the index seeks to represent and measure. For example, the consumer price index (CPI) is a comprehensive measure of changing prices in various economic sectors and a means of monitoring those changes. An index operationalizes a concept by monitoring measurable manifestations that purportedly reflect its essence, presenting cumulative values of the variables measured for each indicator. An index of nominal variables (for example, variables that are list items, with no significance to their order) is basically a typology. While an index usually focuses on one domain, a typology ordinarily examines the intersection between two or more domains. A scale is yet another type of quantification that can be applied to social phenomena. Unlike an index, it usually includes only one domain.

Gender equality can be defined thus: a situation in which there is no difference between women and men in social, economic, and demographic indexes. Today, significant disparities between women and men exist in every variable examined. While examination of each initial indicator is essential, it is also very important to aggregate the indicators into a single gender index that depicts the overall state of inequality between women and men in society.

The Consumer Price Index (A Case Study)

The Consumer Price Index expresses, in a single number, a variety of fluctuations in price that affect a broad range of products and commodities. The conceptual problem with this index is that it does not facilitate clear apprehension of the cumulative changes in prices overall. The simple average of all consumer spending throughout the market, as reflected in the Consumer Price Index, reflects changes in consumer behavior in the designated period, as well as changes in prices. On the other hand, the changes in consumer behavior can be isolated by fixing the quantities of goods and calculating the average prices of these goods, weighted according to the volume of consumption. However, the assignation of weights is random in terms of time, and the question arises as to which year's consumption of each commodity should be measured—since afterward, that year should then remain unchanged to facilitate comparison (that is, that year should become the base year). In fact, all consumer price indexes in the world are based on a fixed basket—that is, a predefined group of products. The basket usually reflects the proportional consumption of products and services in a given period. A fixed basket index can therefore be defined as the ratio between the cost of the basket's contents in one period of time and the cost of the same basket's contents in the base year. Prices are usually compared with those in the base year, when the number of products in the basket was determined.

The methodological rationale behind the CPI can be of use to us in the development of the Gender Index. Gender inequality can be defined as the ratio between inequality manifested in a fixed group of socioeconomic indicators in a given time period and inequality manifested in the same group of indicators in the base period. In keeping with this definition the index can

be expressed in the following formula, which depicts an aggregation of changes occurring in the indicators of gender inequality:

1) $E_t/0^{(c)} = (\Sigma E_t/0*W)/\Sigma W$

E represents inequality in a particular socioeconomic indicator; t represents the period in which it was measured in comparison with the base period (0) in which there was a particular basket of indicators (c). W is the weight assigned to each indicator included in the Index. Σ expresses the sum of the values of all the indicators of gender equality defined by the Index.

E_0 is the inequality manifested in the indicator in question during the base period. E_t is the inequality manifested in the indicator in question during the period being examined.

2) $E_t/0^{(c)} = E_t E_0$

This equation expresses gender equality as represented by the social indicator, relative to the equality represented by the same social indicator in the baseline period.

3) $W = (E_0 Q_c)$[58]

This equation expresses the values of the weights given to each indicator for the purposes of calculating equality, relative to equality in the same indicator in the base year. W is the weight of the indicator in the overall Index based on inequality in this indicator in the baseline period, times its value (Q) during the baseline period (C). As noted above, Σ expresses the sum of the values of all the social indicators included in the group of equality indicators included in the Index. The construct "equality" replaces the construct "cost" used in the Consumer Index.

Methodological Problems in Construction of the Index

When constructing a gender index, it is necessary to determine which indicators are to be included and what weight to assign each one. In the case of the CPI, these issues are resolved using data gathered by periodic surveys conducted by the Central Bureau of Statistics. Such data pertain to the distribution of consumer spending and are presented for a representative sample. In this case the weights express the rate of expenditure on a product from a given basket of goods consumed in the predetermined base period, according to the survey from which the sample is taken. By contrast, the Gender Index has no analogous database that can be used to determine which indicators are to be used and how to assign each of them weights.

58 The significance of the abbreviation Qc rests on the methodological sources of the Gender Inequality Index, which is based on the Consumer Price Index. In the present Index, the "cost" is inequality. The Index is calculated using a formula equivalent to formula (1) above, according to which the index is the cost of products in period t weighted according to consumption in period C, compared with their cost in the base period O. Each cost is multiplied by the consumption rate to make it possible to predict price increases after adjusting for changes in consumption patterns of the various commodities. In order to change the makeup of the basket and disconnect it from the base year, the CPI's products are weighted in such a way that they equal the cost of each product in period C, multiplied by the relative price of this product, which compares its price in the base year O with its price in period C for the period in which the basket was determined.

Edward Harvey, John Blakely, and Lorne Tepperman (1990) propose constructing an index thus: selection and validation of a group of indicators of equality from the array of existing equality indexes, followed by assignation of equal weights for each of the indexes selected. They argue that this approach has the advantage of being based on existing data, making the resulting index easy to compute and inexpensive. It has the disadvantage of having no theoretical or empirical basis for allocating equal weights to all of the indicators. The authors therefore propose a method that makes it possible to measure the weights for each source of inequality at a certain point in time so as to give the weights an empirical basis. The stages they propose for computing the Gender Index are as follows:

- Data selection and preparation
- Data validation
- Data synthesis
- Index calculation

Data Selection and Preparation

This process is analogous to the CPI process of identifying a fixed basket of commodities for a relevant target population. Here we must identify socioeconomic indicators that correlate with the construct of gender inequality. In much the same way that the prices of commodities that are not consumer goods are excluded from the CPI, gender gaps in socioeconomic indicators that do not correlate with the construct of gender equality will be excluded from the Gender Index. The process of data selection and preparation consists of three main steps:

- Face validation: The indicator must be assessed to ensure that it represents what it is intended to represent. The process involves examination of all potential indicators to determine which of them pertain to gender equality. This is achieved by answering this question: Does this indicator actually reflect changes in gender equality as we perceive it?
- Repolarization: This process ensures that all the indicators operate in the same direction so that any growth in equality will be expressed as a growth in the indicator, the ratio, or the rate. For example, the ratio between women's salaries and men's salaries increases when equality prevails. On the other hand, the ratio between the number of women working part time and the number of men working part time decreases as equality increases, and the polarity of the indicator must be reversed to reflect this increase.
- Converting into proportions: This involves converting all indicators into units of rates or percentages (if they are not already so) with a score of between 0 and 1 (or 100%). Calculating the average of indicators expressed in such units is statistically more stable than calculating the average of indicators that express different units.

Data Validation

Validation of the data ensures that indexes of gender equality that we expect to be correlated with one another are indeed correlated, their face validity notwithstanding. The validation method is factor analysis. This procedure requires a given and consecutive series, preferably annual (data collected every year since the base year). The purpose of factor analysis is to

express several variables via a limited number of factors and depict the units of examination in a concise and convenient manner.

The factors are new variables calculated as linear combinations (weighted averages) of the standardized original variables (in other words, each variable has an average of 0 and a variance of 1). Variables require standardization, since different variables are measured in different units: a variable value can be a number, a quotient, or a percentage. It can be measured in New Israeli Shekels or years of schooling. Since our index is an inequality index, all variables are presented in the form of a ratio between the rates of men and women in a given period, thus overcoming the problem of different units. In factor analysis, in order to map the differences in the status of women over time, units are arranged in a series: the first factor is the linear combination that explains most of the variance, and hence discriminates the most between the years. The second factor explains the second largest variance out of the variance that cannot be explained by the first factor, and so on.

In constructing the Index we will use a large group of variables that are by definition highly correlated since they all represent the same social phenomenon. Through factor analysis we can construct a new variable consisting of a series of existing variables. As described above, these factors constitute an orthogonal system of axes within the multidimensional array of variables (because each factor is a linear combination of the original variables and the factors are orthogonal). This kind of analysis is called principal component analysis. The factor analysis also produces the Kaiser-Meyer-Olkin Index of Sampling Adequacy that helps sift through the data. The Kaiser-Meyer-Olkin index has a dual purpose: it can be used to determine whether the variables belong to the same content area, and it can be used to examine a single variable's contribution to the group in which it has been classified. This index's values range from 0 to 1, and therefore results of 0.5 and up indicate that the indicator indeed belongs to the same content area.

Another important concept is factor loading. Factor loadings are coefficients between the variables and the factor. The size of the factor loading measures the relative importance of each variable, by year. A low factor loading on a variable by all factors can be sifted out of the analysis. There are several options in factor analysis, including rotating axes (factors) to increase the association of each variable with one single factor, where possible, and decrease its association with the other factors. This can lead to each factor significantly expressing a different group of original variables that belong to a certain sphere, such as variables that define education level or standard of living. Therefore we must remember that any interpretation of the meanings of factors is but one of many that could have been obtained by rotating the factors.

Factor analysis will generate the first factor, which in this case can be interpreted as the "gender equality factor." We can expect a very high correlation between the component indexes and the first factor; therefore any factor that loads less than 0.71 on this factor will be rejected and excluded. The criterion of 0.71 was selected based on the rationale that at least half of the indicator's variance is explained by the first factor, "the gender equality factor" (namely, variance = loading squared = [0.707] ^2= 0.5). All of the indicators that meet this criterion will be included in the Gender Index. The factor produced will be measured in terms of rate of explained variance and its polarity.

Data Synthesis

At this stage the annual values of the indicators must be standardized. Each value is recalculated relative to its average and the Standard Deviation (SD) of the series. Standardization is the best way to calculate indicators that are originally in different units—such as currency, number of people, or number of hours—and therefore have different distributions and characteristics. Standardization also compares the weights, which are indicative of the relative importance of each indicator in the overall Index. The results of the standardization enable us to compare indicators using the same units. Thus, for example, we could say that between 2000 and 2005, equality improved by an SD of 1 in terms of relative income, while the relative unemployment rate improved by only half an SD—that is, half as much as the income indicator.

An alternative procedure is to use the weight of each indicator in standard scores to influence their relative importance in the overall index. Some, however, claim that in the absence of any logical theory or empirical data that support giving weights to each indicator, they should be evenly weighted (Harvey, Blakely, and Tepperman 1990, 306).

Calculation of the Index

This is the final stage in the calculation of the Gender Index. As we saw from the aforementioned formula, this is when we calculate the weighted average of gender inequality as manifested in the social indicators, relative to the equality situation in the base year. The average in the base year is determined to be 100, and index values are determined accordingly. That is, the standard scores for the base year are transformed to make them equal 100. The standard scores for all the other years are transformed so that their values are less than 100 when equality in an indicator decreases relative to the base year and above 100 when equality increases relative to the base year.

To demonstrate such a transformation: If the standard score of one indicator in the base year is -1.0 (or 1 standard deviation below the indicator's average) we will decide that that value, z = -1.0, equals 100. If the standard score for the following year is -0.25 (or ¼ of a standard deviation below the average), the transformed value will equal the sum of 100 plus the product of the standard deviation and the gap (in terms of standard deviation) between the base year and the following year. Suppose an indicator's standard deviation equals 20. The gap in terms of standard deviation between the base year and the following year is 0.75 (or the gap between -1.0 and -0.25). Therefore, the indicator's value for the next year is 115 (100 + 0.75[20]).

The last stage involves calculating the index using the aforementioned equation for doing so. To illustrate this calculation by way of an example: Suppose we have six indicators in the Gender Index. In the base year each one of them equals 100 and their values for the following year are 110, 115, 120, 105, 100, and 110, respectively. Moreover, all six indicators have equal weights that get the value of 1. Following the equation, the calculation of the index would be as follows:

$$E_t/0(c) = \frac{(110/100)+(115/100)+(120/100)+(105/100)+(100/100)+(110/100)}{1+1+1+1+1+1} = 1.10$$

The accepted practice is to multiply the index by 100 and express it relative to its value in the base year, which would make the result, in this case, 110. The index moved 10 points toward equality between the two periods—from 100 to 110.

Calculating the Gender Index

As indicated above, to calculate the index we need a data series that goes back a few years, preferably an extended one. Indicators usually present two typical problems: some have the necessary number of values but do not have face validity—that is, they do not measure gender equality; other indicators have face validity but do not have enough continuity of measurement or enough values to provide a statistical basis for evaluating missing values.

The Algorithm for Calculating the Gender Index

1. Selection of a series of indicators that have face validity: checking that the indicator assesses what it is intended to measure (face validation). This involves examining all potential indicators to determine which of them is pertinent to the assessment of gender equality. We set out to determine whether each indicator expressed changes in gender equality. This was a two-stage process consisting of descriptive analysis and preliminary screening.

   Descriptive analysis: Analysis of the initial list of variables, including statistical indexes for analyzing the distribution of each one independently according to distribution parameters, distribution symmetry, and extreme values. We also calculated coefficients between pairs of variables. The purpose of these measures was to reduce the number of variables and avoid including overly influential variables or highly correlated variables. When the Pearson coefficient between two variables is higher than 0.8, the possibility of excluding one of them from the Index should be considered. The rule was that variables representing different social phenomena would be included despite being highly correlated, while variables representing the same phenomenon should preferably have symmetrical distribution, high variance (differences between instances), and low correlations between them.

   Preliminary screening: Preliminary screening of suitable indicators was accomplished on the basis of

   a. the indicator's relevance to gender issues in Israeli society;

   b. reliable data for every point of time, to ensure consistency of the Index;

   c. consistency with other studies. It was important to include as many variables used in other studies on gender equality as possible so as to facilitate comparison with other studies and contribute to the discussion of the subject. At this stage we recommend consultation with relevant agents in the field of gender in order to achieve as broad a consensus as possible for the Index.

2. Repolarization of the indicators so that they all express decreasing gender equality: this process ensures that all measures operate in the same direction—that is, increase along with gender equality. For example, the ratio between the salaries of men and women rises when equality is attained. Conversely, the ratio between the part-time employment rates

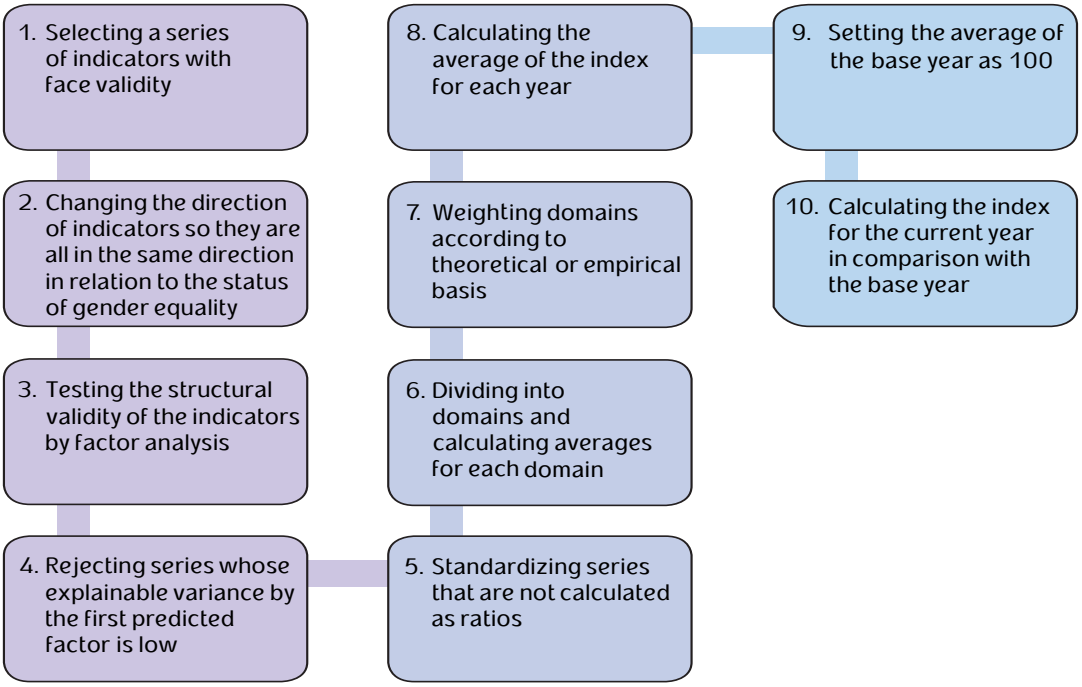
of men and women decreases when equality is achieved, and the direction of this indicator must hence be reversed.

3. Examining the construct validity of the indicators by means of factor analysis: factor analysis is used to predict the gender equality factor that will emerge from the statistical process used to determine what links the variables that have been selected and examines the correlation of each indicator with that factor.
4. Rejecting data sets with an explained variance of less than 0.71 according to the first predictive factor: factor analysis will lead to the factorization of a first factor, which in this case can also be interpreted as the "gender equality factor." We can expect a very high correlation between the component indexes and the first factor; therefore, any factor that loads less than 0.71 on that factor will be rejected and eliminated from the Index. The criterion of 0.71 was selected based on the rationale that at least half of the variance of the indicator (namely, variance = loading squared = $(0.71)^2 = 0.5$) is explained by the first factor, "the gender equality factor." All the indicators that were not rejected in this process will be included in the Gender Index.
5. Standardizing of data as described above.
6. Assigning indicators to domains and calculating the averages in each domain: at this stage we divide the data sets into topics and use factor analysis to reexamine which sets fit together and do not negate each other. The areas in which we will examine gender inequality are naturally those that appear in international gender indexes: the labor market, health, political representation, and so on. Our objective was to challenge the standard list of areas in which gender inequality is measured and identify new ones that are derived from feminist principles. We therefore added the examination of women of different social statuses, violence against women, and women in Arab society in Israel.
7. Weighting by area or on a theoretical empirical basis: at this stage the Index's formula is consolidated and the proper weights are given to each domain. Considerations in determining the various weights might be theoretical (that is, based on a theory regarding which areas influence gender inequality and to what extent) or empirical (based on research data such as a survey of which aspects of inequality bother women more than others). At this stage of the research, we cannot establish weights, and therefore the assumption is that each area has the same impact on the general Index. Accordingly, the formula for calculating the index by its domains is as follows:

   The index = 1/10*(labor market)^2 + 1/10*(violence)^2 + 1/10*(periphery)^2 + 1/10*(Arab society)^2 + 1/10*(poverty)^2 + 1/10*(education)^2 + 1/10*(power)^2 + 1/10*(health)^2 + 1/10*(segregation of professions)^2 + 1/10*(family status)^2

8. Calculating the index average for all the years of measurement.
9. Setting the average for the base year at 100. The base year is randomly determined and serves mainly as a point of origin for comparison.
10. Calculating the current year's Index in comparison with that of the base year. (The base year can be changed and the Index recalculated. The interyear values will be maintained.)

Israel Gender Index Model



Selection of Variables to be Included in the Gender Index

The variables included in the Index must be consecutive series going back several years so their relevance to factors that explain gender inequality can be examined. They must also have potential for continued measurement because our intention is to update the Index every year. Stable data series are mostly found in CBS publications and to a lesser extent in NII publications.

Given that the agencies that have been collecting data in Israel over the years are not committed to gender equality, much of the data collected is not gender disaggregated. Moreover, the categories determined and the subjects researched by those agencies are biased toward the men's sphere and its topical dictates.

There is rarely adequate data regarding women's status and experiences, including violence against women, the distribution of resources within the nuclear family, state allocation of resources by gender, compound disaggregation by ethnicity and gender, and more. These technical but critical constraints limit the indicators that can be included in the Index. The preliminary screening of the relevant variables was performed according to the considerations described above.

Indicators That Were Considered but Not Included in the Present Index

As noted, not all the indicators that were considered were ultimately included in this Index. A list of the excluded indicators and the reasons for their exclusion follows. Some of them may be included in future iterations of the Index.

- Murder of women by their partners: according to the standard analysis—rate per 1000 women. Fluctuation is negligible in this indicator.
- Arrest of men for domestic violence offenses: data are available only for the years 2008–2013.
- Arrest of men for sex offenses: data are available only for the years 2008–2013.
- Prison sentences for men convicted of domestic violence offenses: data are available only for the years 2007–2013.
- Rate of repeated incarceration of prisoners convicted of domestic violence offenses: data are available only for the years 2007–2013.
- Prison sentences for prisoners convicted of sex offenses: data are available only for the years 2007–2013.
- Rate of repeated incarceration of sex offenders: data are available only for the years 2007–2013.
- Rate of participation of men and women aged 15 and up in the civilian labor force in Israel, Ashkenazim versus Mizrahim: does not belong to any existing domain in the index; future indexes will examine ethnic gaps.
- Ratio between the number of Arab women and Arab men who are members of parliament: the number of women Arab MKs is too low.
- Rates of women and men who report diabetes: data were collected only for 2003, 2004, and 2009.
- Rates of women and men who report a disability or a severe disability: data were collected only for 2003, 2004, and 2009.
- Rates of women and men who report physical activity: data were collected only for 2003, 2004, and 2009.
- Rates of women and men with a high BMI: data were collected only for 2003 and 2004.
- Number of domestic violence investigators in the Israeli police by religion and gender: figures were published only for 2010.
- Rate of Jewish versus Arab girls in preschool and public and city day care centers: no gender disaggregated data.
- Division of domestic responsibilities between married people aged 20 and up: data exist only for 2009.
- Overall fertility rate of women in Israel.
- Fertility rate of Arab women in Israel.
- Average age of mother at first birth—overall population.
- Average age of mother at first birth—Arab society.[59]

59 The last four indicators were eliminated because of disagreement over whether they are directly pertinent to gender inequality.

The Indicator Selection Process: Partial Results of Factor Analysis

The statistical procedure of factor analysis allows us to select indicators by determining which of them are associated with the same hypothesized factor. Like many statistical procedures, however, factor analysis requires numerous observations—between 100 and 300 (Field 2005). Another variable that must be taken into account is the number of indicators: the ratio between the number of variables in each indicator and the number of indicators should be at least 1:2. In other words, to examine a list of fifty indicators of gender inequality, each indicator should include at least one hundred observations. Since we are working with years, this is not feasible, and we therefore had to conduct the factor analysis in stages, dividing the indicators into groups (some of the simulations we ran are presented below). Because of these constraints, most gender indexes include strong indicators that have reliable data from consecutive multiannual measurements, but these are not an exhaustive representation of the phenomenon of gender inequality. Since our goal was to present a broader picture of gender inequality, we had to compromise the sample size of each indicator and thereby limit the robustness of the results of the factor analysis used to select indicators.

Before performing the factor analysis, we examined the correlation between each pair of indicators: when no correlation between them was found, we concluded that factor analysis would not provide us with any additional information, because they appeared unrelated to the hypothesized variable of the factor analysis. Conversely, if we found high or full correlation between two indicators, it would be impossible to isolate each of their contributions to the hypothesized factor, which would indicate that the analysis was invalid. Therefore, examining correlations between pairs of indicators was another tool used to select indicators. Another preliminary stage before factor analysis, which also contributed to indicator selection, was an examination of whether their distributions were normal, or at least not overtly contrary to a normal distribution. Sharp ups or downs might have impaired the factor analysis.

Several indicators were thus filtered out—for example, in the areas of education and poverty. All the education indicators are correlated with each other. We selected two—people with 13–15 years of education and people with 16 or more years of education—and the indicators of undergraduate degrees and at least 13 years of education were dropped. Poverty after transfer payments and poverty before transfer payments are highly correlated; we retained the former because it is related to the involvement of the welfare system. Because of the limitations of the analysis—a result of the limited number of observations for each indicator—we performed factor analysis of each indicator within its domain and examined the correlations of each indicator with other similar indicators and with all indicators. This constituted another stage of indicator selection. Table II-a below shows the results of the factor analysis of the variables.

Table II-a

Factor Analysis of All Indicators in the Index

Component Matrix*

		Component							
		1	2	3	4	5	6	7	8
w_participation100	labor market participation	.900							
w_parttime100	Part-time work	.626	.520						
w_salary100	Average monthly salary	.907							
w_wage100	Average hourly wage		.654				.601		
w_temp100	Contract workers						-.519	.508	
w_housewife_part	Housewives working part-time	-.824							
w_housewife_unempl	Unemployed housewives			-.495	-.529	.432			
w_median_salary	Median monthly salary		-.792	.474					
educ_13_15_100	Those with 13-15 years of education	.438	.731						
educ_17_100	Those with 16+ years of education	.804	-.501						
pov_after100	Prevalence of poverty		-.739						
pov_benef100	Income support	.584	.599						
pol_parl100	Members of parliament	.828							
pol_minister100	Government ministers	-.800							
pol_muni100	Mayors/Council Heads				.662		.437		
pol_general_manager100	CEOs	.823				.512			
pol_senior_manager100	Senior managers	.734	-.630						
pol_other_manager100	Other managers	.821							
pol_highest_position100	Senior civil servants	.973							
pol_senior_contract100	Senior contracted civil servants	.890							
h_expectancy100	Life expectancy	-.825							
h_mortality100	Mortality rate		.713					-.433	
h_feeling100	Perception of health	.541	-.742						
a_participation100	Arab society labor market participation	.965							
a_parttime100	Arab society part-time		.722						
a_salary100	Arab society average monthly salary	.702		-.640					
a_wage100	Arab society average hourly wage				.496	-.651			
a_educ1315_100	Those with 13-15 years of education – Arab society	-.581		.410	.648				
a_educ16_100	Those with 16+ years of education – Arab society	.895							
a_complaints100	Domestic violence complaints by Arab women	-.912							
a_fertility15_19	Arab teen pregnancies	.956							
a_age_marry	Arab society marriage age					.719			
v_sex_abuse_complaints100	Calls to rape crisis centers	.862							
v_sex_abuse_cases100	Cases transferred to police	-.659			-.514				
v_violence_abuse_victims100	Women treated at welfare centers	-.903							
v_violence_cases100	Domestic violence cases	.826			.444				
v_violence_cases_closed100	Cases closed for lack of evidence			-.504	.416	.425			
p_participation100	Participation in the periphery	.674			-.690				
p_salary100	Monthly salary in the periphery			.789					
seg_engineer_architect100	Engineers and architects	.736		.507					
seg_doctors100	Doctors and pharmacists	.683		-.655					
seg_judge_lawyer100	Judges and lawyers		.701			.424			
seg_teaching100	Educators	.781		.498					
seg_carry100	Caregivers			.772		.429			
seg_hitec100h	HiTech employees	.758	.424			-.456			
segregation_professions	Segregation of professions	.784	.535						
segregation_disciplines	Segregation of occupations	.754	.515						
family_fertility_15_19	Teen pregnancies	.939							
family_single_parent_family	Single parent families	.480	.585						
family_age_marry	Average marriage age	.533	-.638						

Extraction Method: Principal Component Analysis.

* 8 components extracted.

Table II-b

Component Matrix*

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	22.910	45.819	45.819
2	8.425	16.850	62.669
3	5.664	11.328	73.997
4	4.418	8.835	82.832
5	3.432	6.863	89.695
6	2.109	4.218	93.913
7	1.888	3.775	97.688
8	1.156	2.312	100.000

As Table II-b shows, 46% of the variance of all indicators over the 2004–2012 period is explained by the first factor, 17% by the second factor, and 11% by the third factor. Table II-c below presents several samples of factor analysis on various domains of the Index.

Table II-c

Samples of Domain Factor Analyses

Labor Market Domain Indicators

Component Matrix*

	Component			
	1	2	3	4
w_participation100	.826			
w_parttime100	.789			
w_salary100	.973			
w_wage100	.488	.504	-.677	
w_temp100		.608	.634	
w_housewife_part	-.726	.498		
w_housewife_unempl			.503	.781
w_median_salary		-.908		

Extraction Method: Principal Component Analysis.

* 4 components extracted.

Power Domain Indicators

Component Matrix

	Component	
	1	2
pol_parl100	.827	
pol_minister100	-.824	
pol_muni100	-.448	.807
pol_general_manager100	.888	
pol_senior_manager100	.833	
pol_other_manager100	.883	
pol_highest_position100	.956	
pol_senior_contract100	.810	

Extraction Method: Principal Component Analysis

* 2 components extracted.

Health Domain Indicators

Component Matrix

	Component
	1
h_expectancy100	.875
h_mortality100	.919
h_feeling100	-.888

Extraction Method: Principal Component Analysis

* 1 component extracted.

Arab Society Domain Indicators

Component Matrix*

	Component			
	1	2	3	4
a_participation100	.953			
a_parttime100	.569			.422
a_salary100	.802		-.579	
a_wage100		.756	-.554	
a_educ1315_100	-.591	.554	.417	
a_educ16_100	.831		.450	
a_complaints100	-.986			
a_fertility15_19	.938			
a_age_marry		-.500		.750

Extraction Method: Principal Component Analysis

* 4 components extracted.

Violence Domain Indicators

	Component	
	1	2
v_sex_abuse_complaints100	-.743	
v_sex_abuse_cases100	.851	
v_violence_abuse_victims100	.922	
v_violence_cases100	-.923	
v_violence_cases_closed100	.514	.856

Extraction Method: Principal Component Analysis

* 2 components extracted.

Gendered Segregation of Professions Domain Indicators

Component Matrix*

	Component		
	1	2	3
seg_engineer_architect100	.878		
seg_doctors100		-.800	
seg_judge_lawyer100		.651	.602
seg_teaching100	.797		-.547
seg_carry100		.877	
seg_hitec100h	.810		
segregation_professions	.911		
segregation_disciplines	.899		

Extraction Method: Principal Component Analysis

* 3 components extracted.

Family Status Domain Indicators

Component Matrix*

	Component	
	1	2
family_fertility_15_19	.915	
family_single_parent_family	.748	-.568
family_age_marry		.906

Extraction Method: Principal Component Analysis

* 2 components extracted.

The Index without Each of the Domains

This appendix addresses the question of the degree to which the Index is dependent on each one of its constituent domains. To this end we examined the trajectory of the Index with each one of the domains removed in turn. The result was as anticipated: the Index is not dependent on any particular domain. Nevertheless, although we grouped them, the units in each indicator and domain are different, and the degree of variance in each indicator is hence derived from the size of the population in which it was measured. For example, in the political power domain, the rate of change is high and the fluctuations more dramatic because the numbers of parliament members and ministers are low. On the other hand, the influence of one woman in a position of power is much larger than that of another woman who is, say, an active participant in the labor market. We therefore contend that changes in the numeric ratio between women and men in politics are likely to have great impact on the Gender Index. The table below presents the impact of each domain on the overall Index, and Figures 3a-j present the Index with each one of the domains removed in turn. They show that in no case does the removal of a domain affect the Index's trajectory; it remains stable in all cases (when the Arab society and family status domains are removed, the incline is somewhat less steep, but the trend remains the same; the removal of the poverty domain has a negligible effect; the removal of the power domain makes the incline steeper but does not affect the trend; the removal of the other domains has no effect whatsoever). Figure 3k is a repetition of Figure 1 in the report itself, presenting the overall Index with all ten domains.

Discrepancy between the Overall Index's Incline and the Incline When Each One of the Domains Is Removed

Removed Domain	Index Formula	Discrepancy between the Overall Index's Incline and the Incline When Each One of the Domains Is Removed
Labor market	y = -0.0065x + 1.0437	0.001
Segregation of professions	y = -0.0045x + 1.0438	-0.001
Violence against women	y = -0.007x + 1.0396	0.0015
Periphery	y = -0.0056x + 1.0461	0.0001
Arab society	y = -0.0043x + 1.0448	-0.0012
Poverty	y = -0.0055x + 1.044	0
Education	y = -0.0051x + 1.0471	-0.0004
Power	y = -0.0068x + 1.0306	0.0013
Health	y = -0.0062x + 1.0463	0.0007
Family status	y = -0.0033x + 1.0429	-0.0022

As the table shows, without the Arab society and family status domains, the slope becomes a little bit smaller but the overall trend remains. The poverty domain has negligible effect, and without the power domain the slope is steeper, but even without it, the Index's trajectory remains the same. Other domains have no impact.

Figure 3a

The Gender Index without the Labor Market Domain

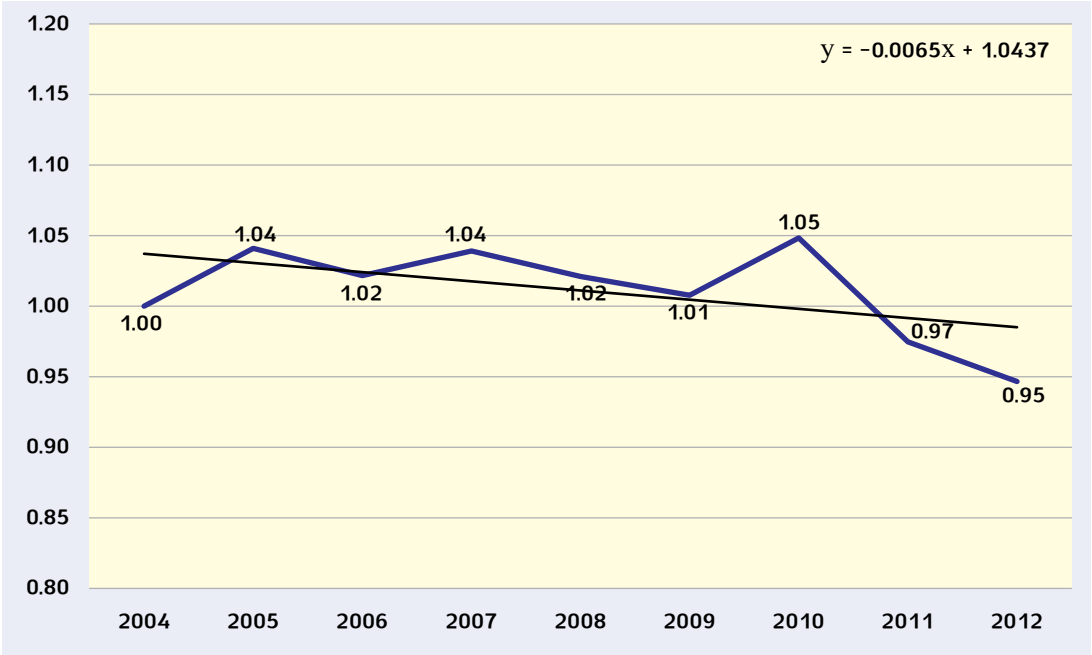


Figure 3b

The Gender Index without the Gendered Segregation of Professions Domain

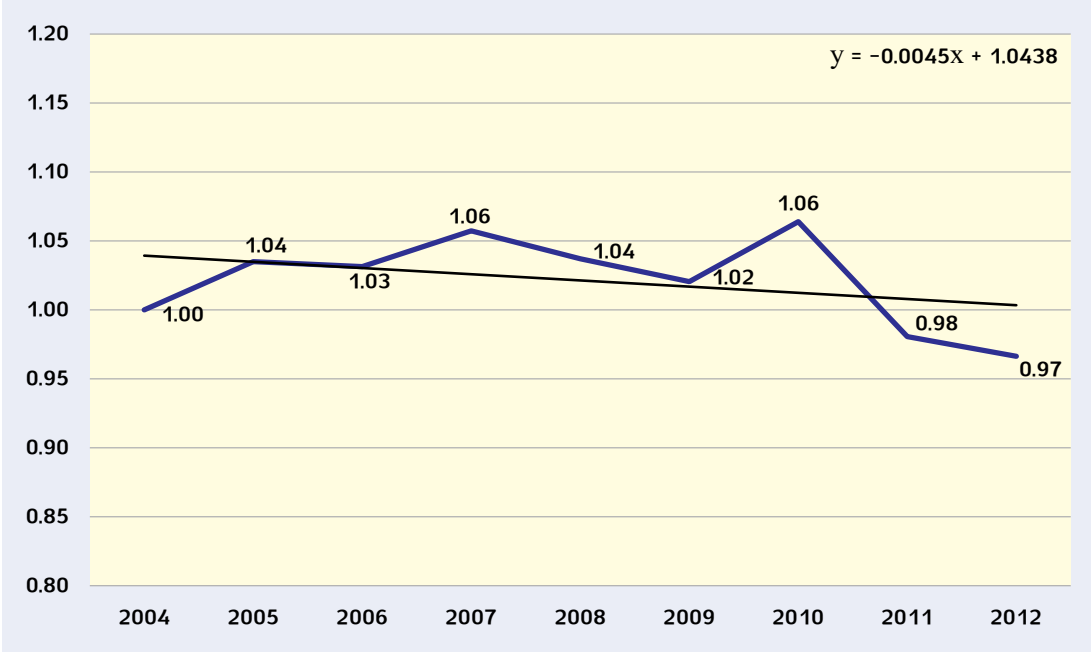


Figure 3c

The Gender Index without the Violence against Women Domain

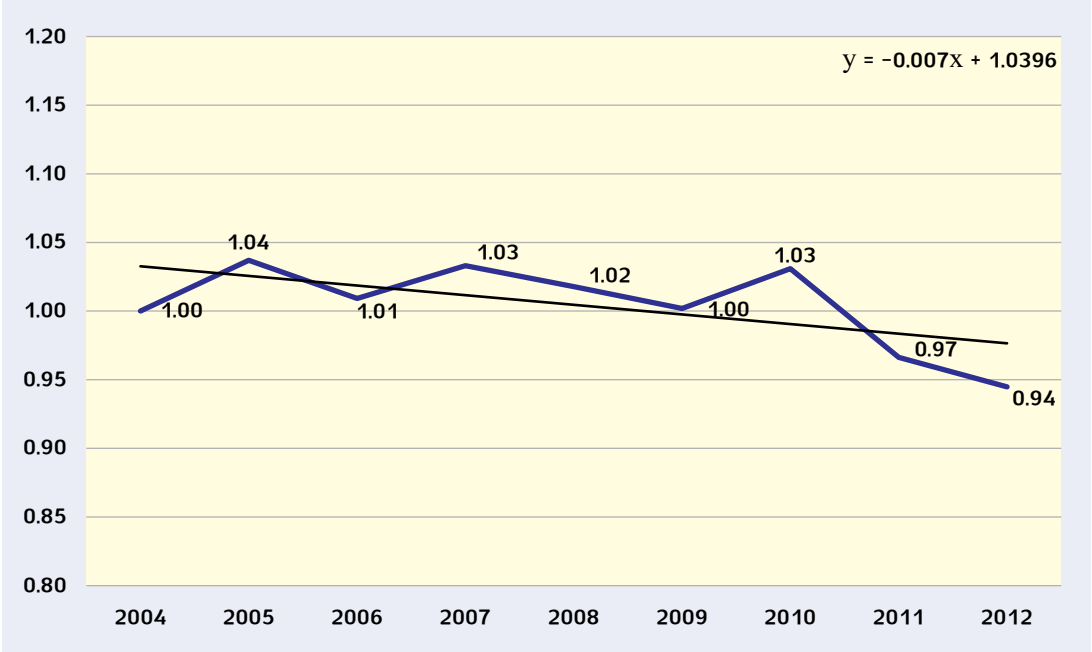


Figure 3d

The Gender Index without the Periphery Domain

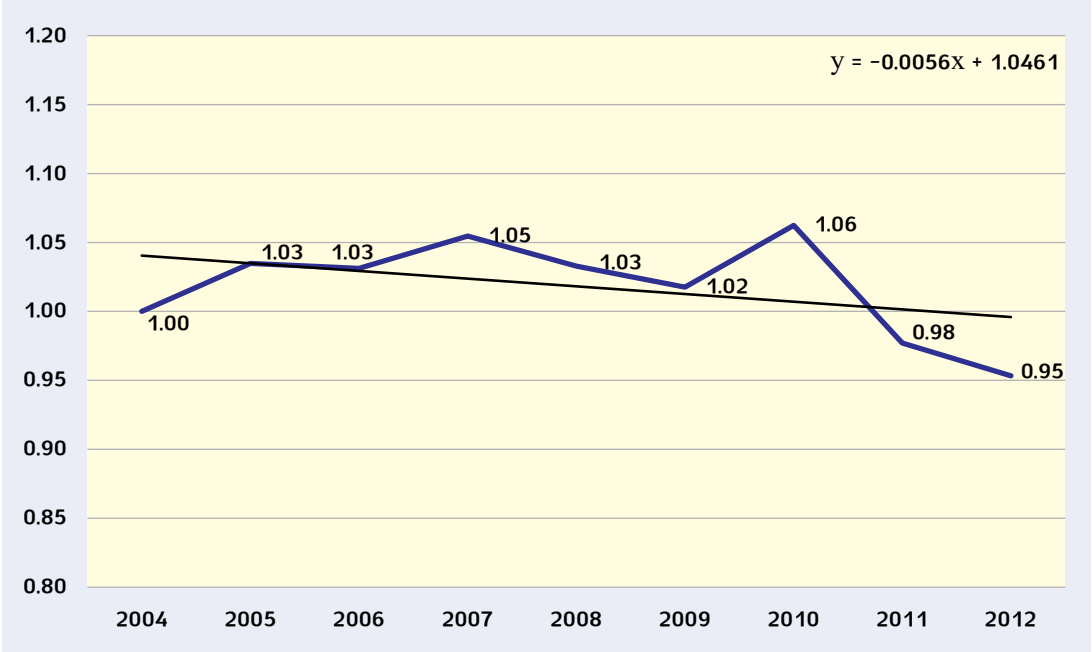


Figure 3e

The Gender Index without the Arab Society Domain

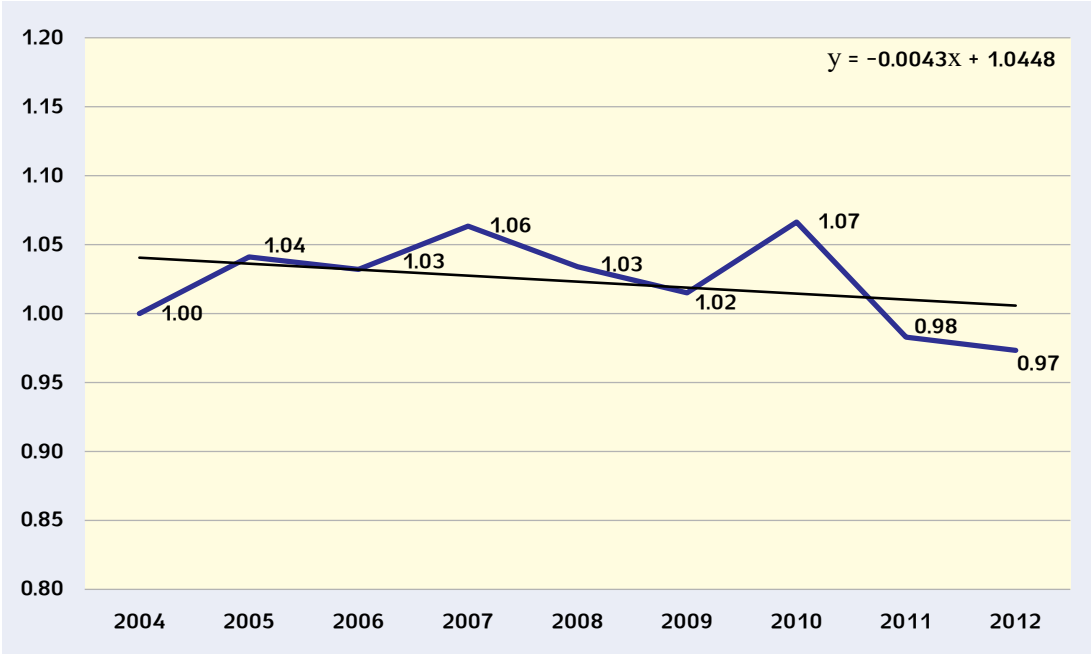


Figure 3f

The Gender Index without the Poverty Domain

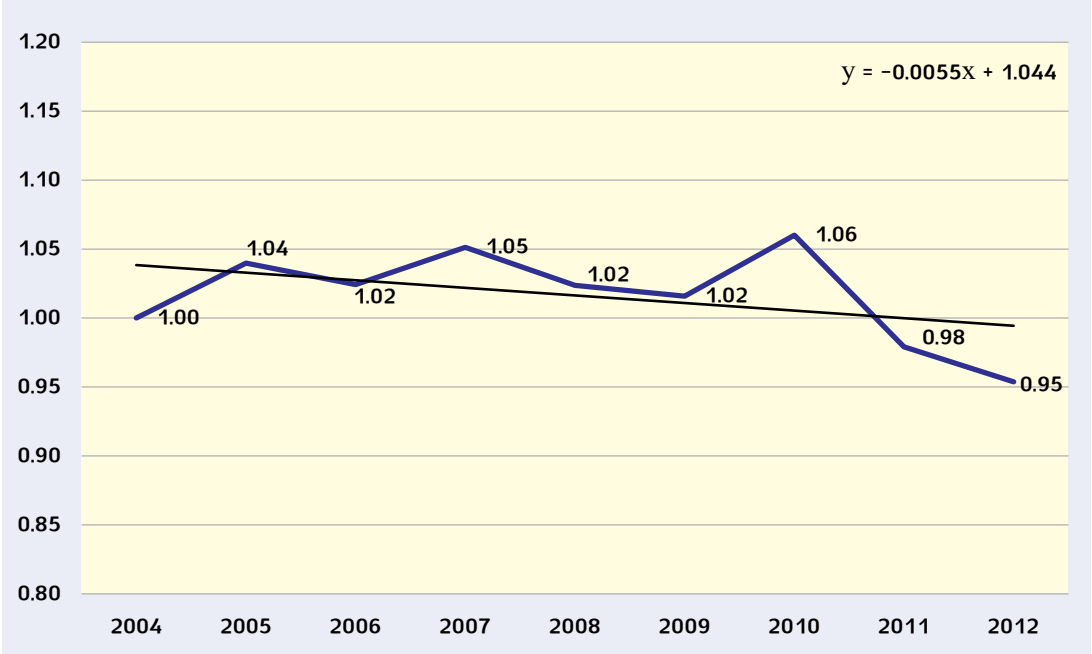


Figure 3g

The Gender Index without the Education Domain

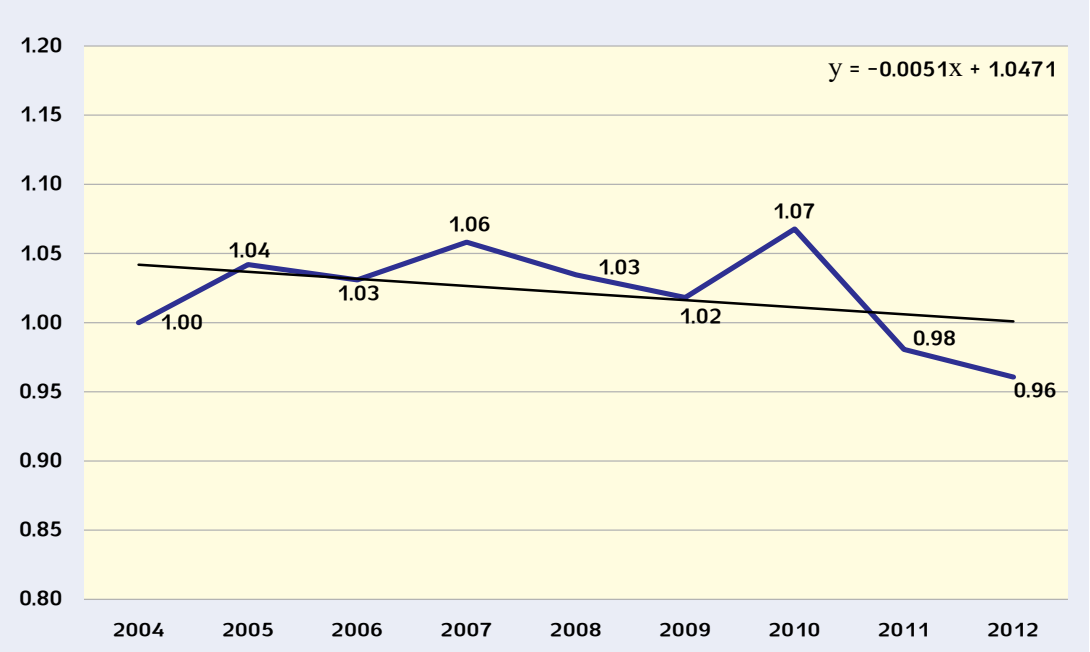


Figure 3h

The Gender Index without the Power Domain

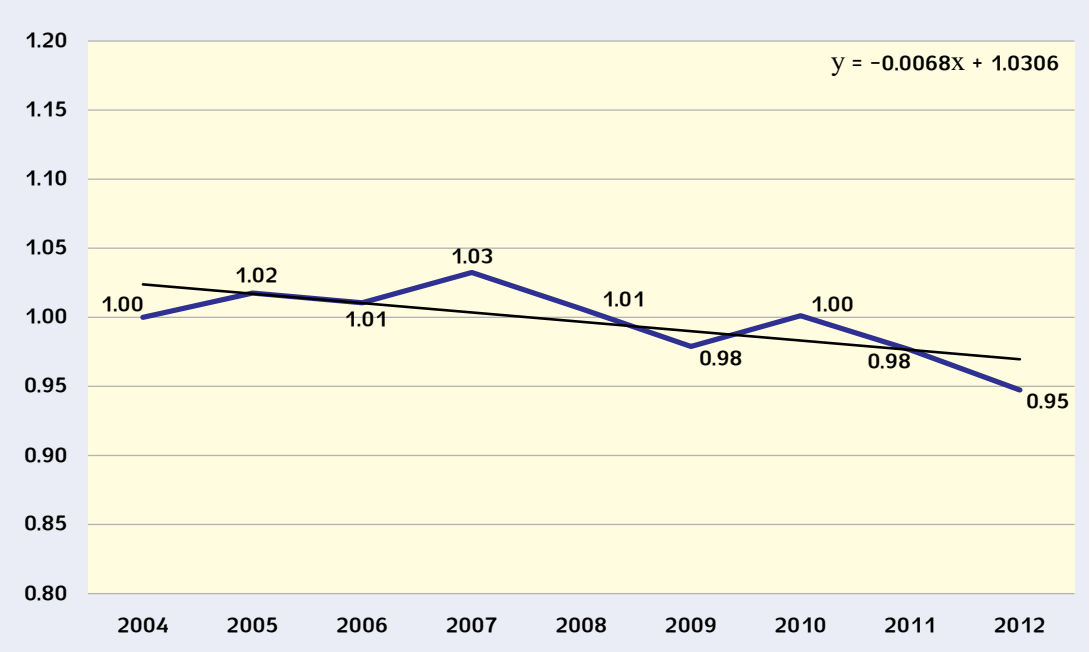


Figure 3i

The Gender Index without the Health Domain

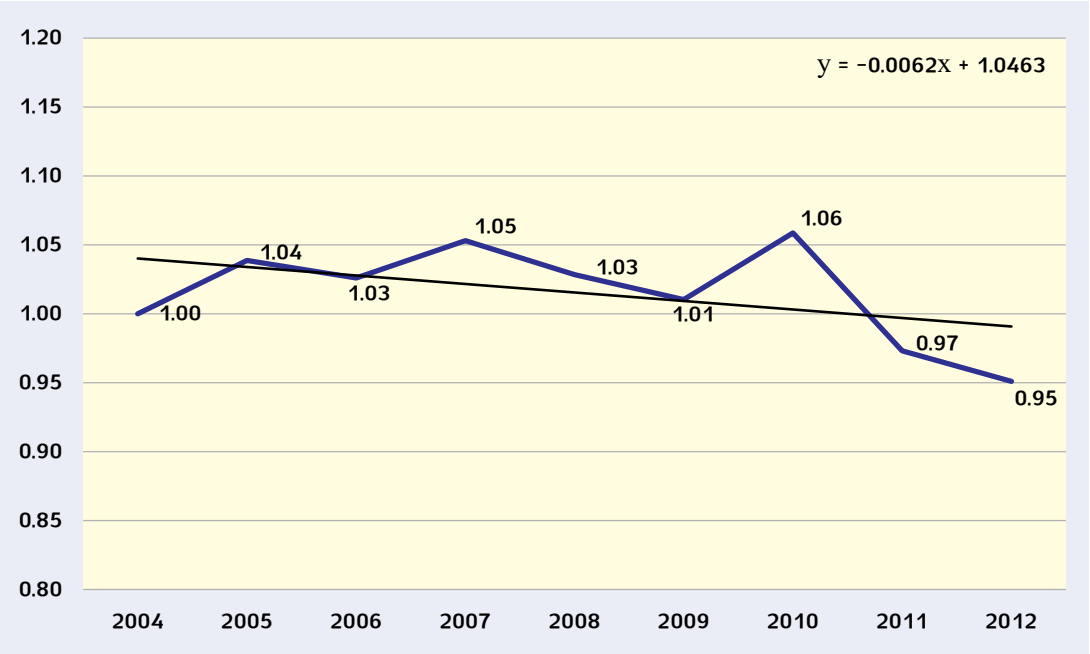


Figure 3j

The Gender Index without the Family Status Domain

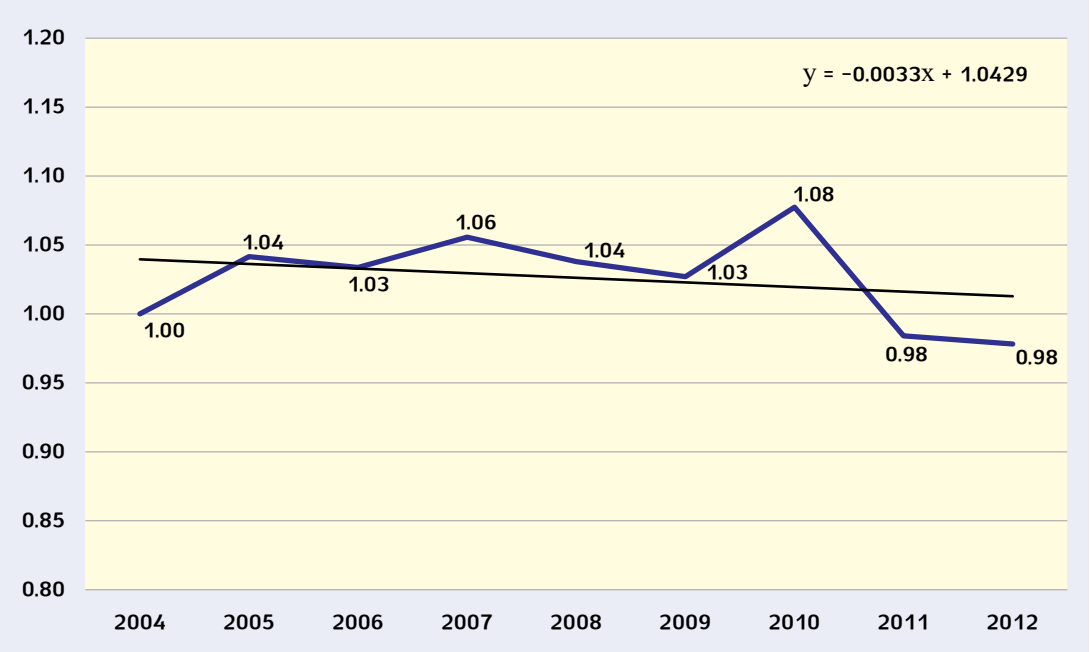
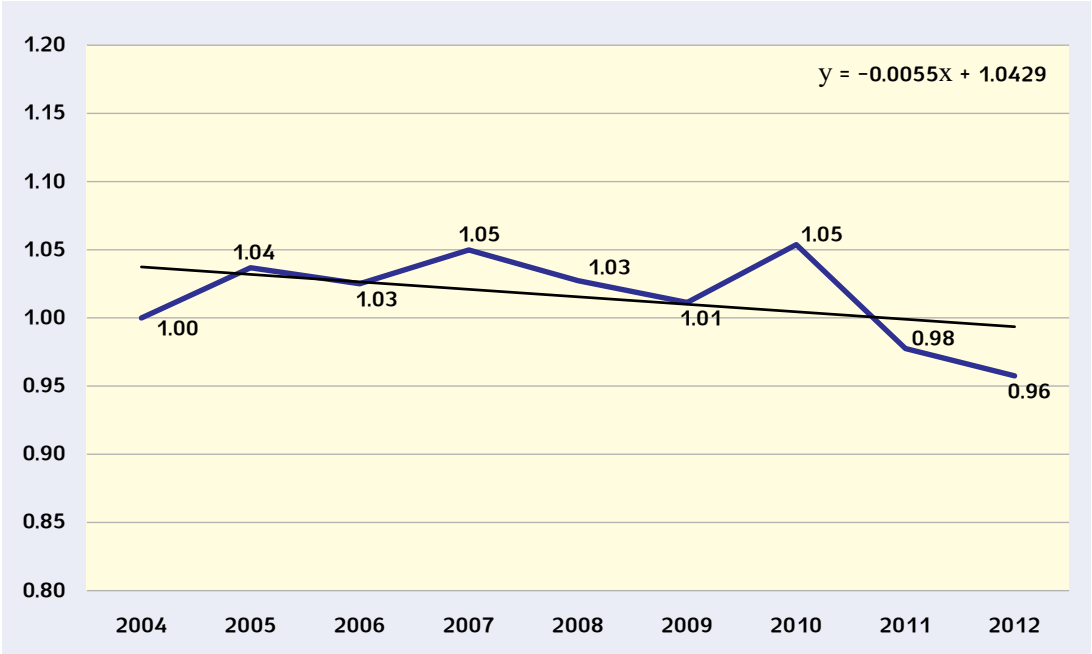


Figure 3k

The Gender Index with All Ten Domains



The Index without Squaring of the Domains

As noted above, the formula for calculation of this Index is:

Index = 1/10*(labor market)^2 + 1/10*(violence)^2 + 1/10*(periphery)^2 + 1/10*(Arab society)^2 + 1/10*(poverty)^2 + 1/10*(education)^2 + 1/10*(power)^2 + 1/10*(health)^2 + 1/10*(professional segregation)^2 + 1/10*(family status)^2

This formula was employed by the OECD in the Social Institutions and Gender Index (SIGI) to measure gender inequality in social institutions and to connect the various domains. The formula assigns equal weights to all the domains. As noted, in the absence of an empirical or theoretical means of determining the weights, since we have no way of prioritizing one domain over another, the common solution is equal distribution. The function selected by the OECD is not linear, assuming that gender inequality is connected to the disadvantaging of women. When inequality increases, this disadvantaging increases at a higher rate, and inequality has more weight in each domain. Nonlinearity also implies that there can be only partial compensation between the Index's domains. In other words, high inequality in one domain can only partially be counteracted by low inequality in another domain.

These considerations led us too to adopt this formula for estimating inequality. However, we wanted to examine the impact of the choice to square each domain. To this end we examined the results of the Index without squaring each domain by means of an equation that aggregates them all. Figure 4a shows the rates of change from year to year without squaring. Figure 4b presents the results of the Gender Index with equal weights for each domain without squaring of their values.

Figure 4a

The Gender Index 2004–2012: Rates of Change in Each Year in Comparison to the Preceding Year, without Squaring

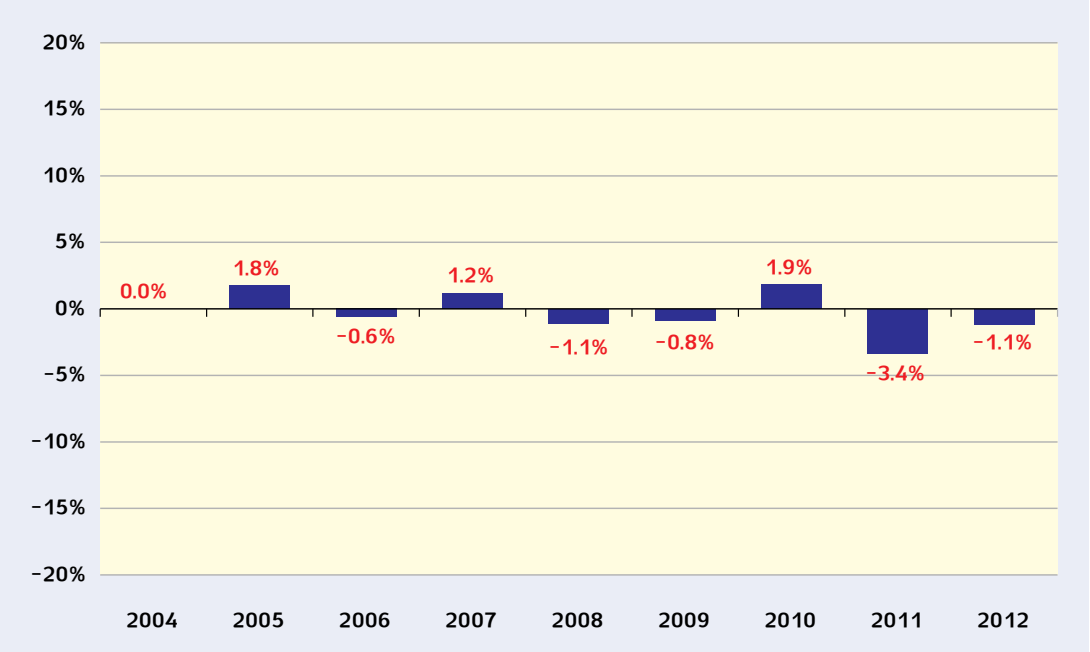
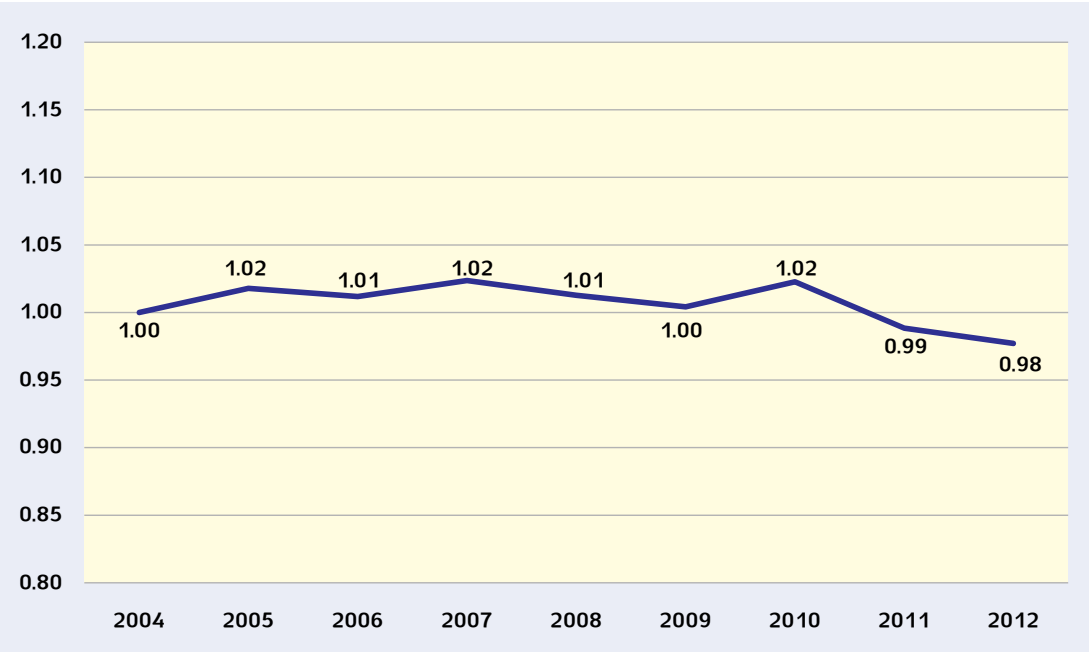


Figure 4b

Results of the Gender Index 2004–2012, without Squaring



The figures show that without squaring, inequality is more stable and its highs and lows less extreme, but the trend remains the same. A small increase in inequality in 2004–2007 is evident. Between 2008 and 2009 inequality decreased, but in 2010 it increased, and in 2011–2012 there was improvement and inequality decreased.