J Popul Econ (2004) 17:45–65
DOI 10.1007/s00148-003-0180-z

Journal of
Population Economics

Did the association between fertility and female employment within OECD countries really change its sign?

Tomas Kögel

Max Planck Institute for Demographic Research, Konrad-Zuse-Strasse 1,
D-18057 Rostock, Germany (Fax: +49-381-2081-446; e-mail: Koegel@demogr.mpg.de)

Received: 21 November 2001/Accepted: 22 April 2003

Abstract. Recent literature finds that in OECD countries the cross-country correlation between the total fertility rate and the female labor force participation rate, which until the beginning of the 1980s had a negative value, has since acquired a positive value. This result is (explicitly or implicitly) often interpreted as evidence for a changing sign in the time-series association between fertility and female employment within OECD countries. This paper shows that the time-series association between fertility and female employment does *not* demonstrate a change in sign. Instead, the reversal in the sign of the cross-country correlation is most likely due to a combination of two elements: First, the presence of unmeasured country-specific factors and, second, country-heterogeneity in the magnitude of the negative time-series association between fertility and female employment. However, the paper does find evidence for a reduction in the negative time-series association between fertility and female employment after about 1985.

JEL classification: J10, J11, J13

Key words: Total fertility rate, female labor force participation rate, econometrics of panel data

---

I benefited from stimulating discussions with Arnstein Aassve, Pau Baizan, Francesco Billari, Henriette Engelhardt, Hans-Peter Kohler and Alexia Prskawetz and a seminar in Rostock. Further, I am grateful to two anonymous referees for very useful suggestions that improved essentially the content of the paper. In addition, I thank Susan Masur, Susann Backer, and Elizabeth Zach for language editing. The views expressed in this paper are the author's own views and do not necessarily represent those of the Max Planck Institute for Demographic Research.
Responsible editor: Junsen Zhang.

1. Introduction

Some salient aspects of contemporary, advanced societies are below-replacement fertility rates and increased female participation in the labor force. Many researchers believe that these features are related. They reason that childrearing and female employment are incompatible, often forcing women to manage under increasing time constraints. At the same time, childrearing remains primarily the responsibility of women (rather than of men). Most OECD countries have a completely or partially unfunded pension system (that is, a pension system in which the current working generation must finance the pension benefits of the previous working generation). Low fertility rates reduce the potential sustainability of this type of pension system. In contrast, high female labor force participation increases its sustainability. (Both factors affect the number of workers who can contribute to pension benefits). Hence, an understanding of the relationship between fertility and female employment at the macro-level is relevant and important to current policy-making.

Most population economics studies dealing with micro-level data show that female wages in real terms and female education have a negative effect on fertility and a positive effect on female employment. This finding implies a negative (and not strictly causal) association between fertility and female employment. Furthermore, most micro-level studies in demographic literature confirm a negative association between these two variables. As I will be discussing in the next section, a recent analysis reveals the existence of a different pattern at the macro-level. The analysis shows that between OECD countries, the cross-country correlation between the total fertility rate (TFR) and the female labor force participation rate (FLP) has changed from a negative value until the beginning of the 1980s to a positive value today. Rindfuss et al. (2004) and Brewster and Rindfuss (2000) explain this reversal with policies that minimize incompatibilities between childrearing and female employment. This new macro-level evidence challenges previous findings, and could be very good news for policy makers. If correct, then such policies imply that a rising FLP increases the TFR. It goes without saying that this would significantly improve the prospects for sustaining OECD pension systems. (Of course, a positive association between the TFR and the FLP could also be interpreted as very bad news for policy makers, since other policies imply that a falling TFR reduces the FLP, which would reduce prospects for sustaining OECD pension systems).

This paper, however, moderates to some extent the optimistic viewpoint just mentioned. In contrast to the aforementioned earlier literature, it uses panel data techniques to pooled cross-country and time-series data from OECD countries. These methods identify and account for unmeasured country-specific factors (henceforth country effects). With these methods, I show that in the time-series dimension within countries, there was *not* a change in sign for the association between the TFR and the FLP. On the other hand, the present study does find support for a falling magnitude and significance of the negative time-series association after 1985. Due to some incompatibility between childrearing and female employment, a rising FLP has indirectly a negative effect on the future pension system by reducing fertility. Therefore, the finding of a falling magnitude of the negative time-series association between the TFR and the FLP is still relatively good news for policy makers. It could mean that policies that minimize incompatibilities between childrearing and female employment reduce this indirect effect of a rising FLP on the future pension system.

Furthermore, the presence of country effects implies that possibly cross-country differences in public policies or labor market institutions might have caused high fertility and high female employment in some countries and low fertility and low female employment in other countries. Galor and Weil (1996) present a general equilibrium model with an endogenously rising relative wage of women, endogenous fertility and endogenous female employment. Their mechanism suggests that the introduction of child care services along with an increasing relative wage of women may generate a positive association between the TFR and the FLP. Recent work presents formal models that introduce child care services into the model of Galor and Weil. Apps and Rees (2001) show that in this framework increasing child care subsidies (or, alternatively, cross-country differences in child care subsidies) can produce a positive association between the TFR and the FLP. Martinez and Iza (2004, this issue of the *journal*) show that one can generate in this framework such a positive association with a rising relative wage of skilled labor (or, alternatively, cross-country differences in the relative wage of skilled labor). However, the results in this paper show that changes in public policies or labor market developments cannot have caused that a rising FLP increases the TFR within countries over time.

Section 2 briefly surveys recent literature that found a changing sign in the association between the TFR and the FLP in cross-country data. The section also explains the motivation underlying the econometric approach applied in this paper. Section 3 presents the aforementioned panel data techniques applied to pooled cross-country and time-series data of OECD countries. Section 4 presents results where, in addition to accounting for country effects, the time-series association between the TFR and the FLP is also allowed to be heterogeneous between three broad country groups (namely, Scandinavian countries, Mediterranean countries and the remaining countries of the OECD). Finally, Sect. 5 contains a conclusion.

2. Motivation

Ahn and Mira (2002) and Rindfuss et al. (2004) recently showed that the annual cross-country correlation coefficient between the total fertility rate (TFR) and the female labor force participation rate (FLP) in OECD countries had changed its value from a negative value (around 1985) to a positive one (see also Benjamin 2001). Figure 1 replicates their results for twenty-one OECD countries for 1960–1999 (all countries and data sources of the figures and tables in this paper are shown in Appendix D).

Further, Brewster and Rindfuss (2000) and Esping-Andersen (1999) showed that in an OLS regression, with cross-country data of the OECD with the TFR as the dependent variable and the FLP as the independent variable, the coefficient of the FLP was significant and negative in the 1970s but significantly positive by the 1990s.[1]

Contrary to this finding, Engelhardt et al. (2004) found in macro-level time-series data from six representative OECD countries that, for all of these countries, the value of the time-series association between the TFR and the FLP did not change from negative to positive.[2] However, in all

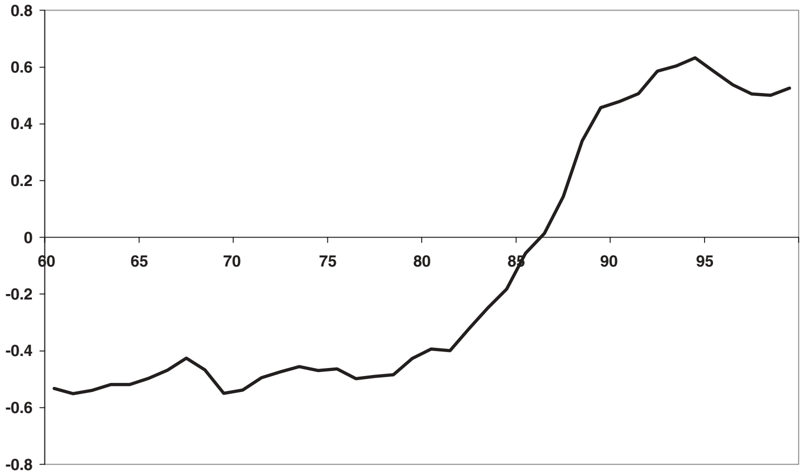

Fig. 1. Annual cross-country correlation coefficient between TFR and FLP in OECD countries

non-Mediterranean countries in their study, the time-series association became significantly weaker and less significant after about the mid-1970s.

Ahn and Mira (2002) argue that the income effects of female wage increases, high unemployment in Mediterranean countries, and extensions of standard economic theory, such as discrete working hours and purchased child-care, could explain the change in sign of the cross-country association between the TFR and the FLP. Somewhat differently, Rindfuss et al. (2004), and Brewster and Rindfuss (2000), argue that changes in the institutional context, such as changing social norms toward working mothers, evolving family policies (such as, cash benefits, and increasing child-care availability), all reduced the incompatibility between childrearing and female employment.

Figure 2 illustrates a hypothesis, which motivated me to re-examine the evidence using panel data techniques.[3] In this figure, I plot the TFR for Italy and Sweden in 1965 and 1995 on the y-axis and the FLP for these countries and years on the x-axis. If one is willing to accept that these two countries are representative of the OECD countries, then the figure illustrates that the reversal in the sign of the cross-country association between the TFR and the FLP is due to a combination of two elements. First, there are country effects, which cause in both years the FLP to be higher in Sweden than in Italy.[4] Second, the negative time-series association between fertility and female employment is weaker for Sweden than for Italy. The figure shows that both elements together imply a changing sign in the cross-country association between the TFR and the FLP, while for each country the association in the time dimension is negative.

Of course, it could be that Italy and Sweden are not representative of the OECD countries. In that case it could be that other factors than country effects and heterogeneity in the time-series association contributed to the reversal in the sign of the cross-country association. However, in Sect. 4 the paper finds empirical support for country effects and heterogeneity in the time-series association between the TFR and the FLP for the broad country groups of Scandinavian, Mediterranean and the remaining countries. Hence, it is at least very likely that these two elements caused

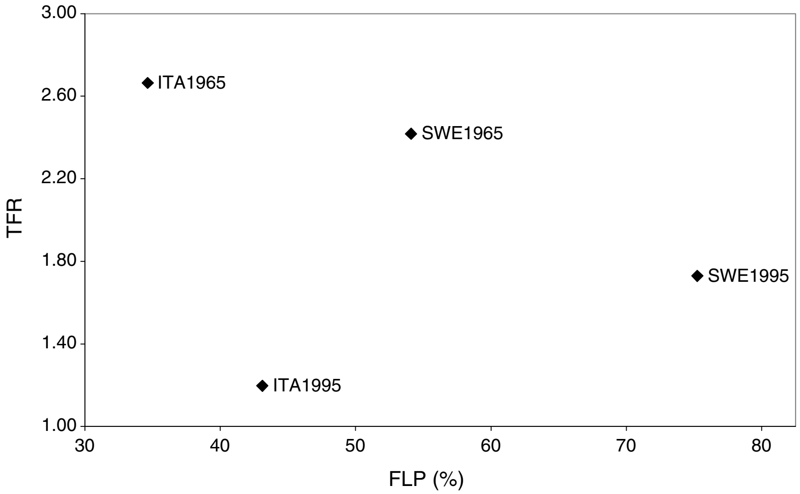

Fig. 2. TFR and FLP in Italy and Sweden in 1965 and 1995

together the reversal in the sign of the cross-country association between the TFR and the FLP.

Note, that in Fig. 2, in the aforementioned literature, and in the next two sections in this paper, the FLP always refers to the female labor force participation of women aged 15 to 64. However, women above age 44 rarely bear children and are often no longer involved in childrearing (at least of very young children). Appendix B shows empirical results with the FLP of women aged 25 to 44. It turns out that the results are almost completely unaffected from this choice of the FLP definition. Instead, the hypothesis of this paper is reinforced with data of the FLP of women aged 25 to 44.

The study by Engelhardt et al. (2004), which uses time-series data from six developed countries, used "cointegration" techniques and found that the TFR and the FLP are causally related in both directions. This finding is consistent with the view that both variables are simultaneously influenced by common exogenous variables such as wages, institutions, and social norms. Earlier literature that examines micro-level data found conflicting results on the direction of causality (see, Cramer 1980; Lehrer and Nerlove 1986). To allow for the possibility of causality in both directions, I applied estimations where the TFR is regressed on the FLP, as well as estimations where the FLP is regressed on the TFR. This procedure is, in principle, compatible with economic theory, which views the TFR and the FLP as endogenous variables that are influenced simultaneously by the real female wage and other primarily economic variables.

Due to their simultaneity, regressing the TFR on the FLP (or the other way around) is not very satisfactory from the point of view of standard economic theory. Instead, it would be preferable to regress the TFR and the FLP on real female and male wages in separate equations. If, after the beginning of the 1980s, the effects of the female wage in real terms on the TFR and the FLP have opposite sign, then the hypothesis of a positive association between the TFR and the FLP is rejected. Otherwise, it is accepted. However, other external variables such as rising social acceptance of working mothers – for which almost no macro-level data exist – seem at least equally important for long-term trends. For this reason, the variable FLP (and TFR) might contain more information than real wages. As a consequence, regressing the TFR on the FLP (and the other way around) seems more appropriate for the interdisciplinary literature that discusses this changing association in cross-country data – a literature that often questions the importance of real wages for fertility.

3. Accounting for unmeasured country-specific factors

In the last section, I argued that the combination of country effects and country-heterogeneity in the magnitude of the negative time-series association between the TFR and the FLP very likely caused the cross-country correlation to reverse its sign. To test the plausibilty of this hypothesis more formally, this section applies econometric methods that can detect and control for country effects, while also assuming homogeneity in the magnitude of the time-series association between the TFR and the FLP. The next section applies the same econometric methods, but also allows for heterogeneity in the time-series association between these two variables for three broad country groups (Scandinavian countries, Mediterranean countries and the remaining countries).

The presence of a time-series dimension in the data is a prerequisite for using econometric methods that account for country effects. For this purpose, the data of Fig. 1 (which were used to calculate the annual correlation coefficient for twenty-one OECD countries) were pooled to a single data set and used for joint estimation. However, I used only quinquennial data (i.e., only the data points 1960, 1965, 1970, 1975, 1980, 1985, 1990, 1995 and 2000) because they are less likely to be serially correlated than annual data. Moreover, to test whether the sign of the time-series association between the TFR and the FLP was negative before 1985 and positive afterwards (as might seem to be the case according to Fig. 1), I divided the data set into two sub-samples: 1960–1985 and 1985–2000.

As is standard in most applied macro-econometric work, all variables in the following regressions were included in natural logarithms. It is possible that the variables in this study are difference-stationary, in other words, the mean and variance is constant over time after first differencing, but not in levels. If true, this could give rise to a spurious regression problem (Granger and Newbold 1974). Appendix C contains panel data unit tests applied to the variables in this study. They have critical values taken from Harris and Tzavalis (1999). It turned out that difference-stationarity could be rejected for all variables in this study. The reason for this is most likely the fact that the data have a smaller time dimension as compared to the cross-section dimension. The result of the unit root tests implies that it is appropriate to apply standard inference to the estimation results.

In the following, only regression results with the TFR as the dependent variable and the FLP as the independent variable are shown, because the results from regressions of the FLP on the TFR are very similar.

Table 1 shows the estimation results for the sub-sample 1960–1985. The second column in this table shows the results of between-group estimation. Between-group estimation is a regression of $\overline{\ln(TFR_i)}$ on $\overline{\ln(FLP_i)}$, where

$\overline{\ln(TFR_i)}$ and $\overline{\ln(FLP_i)}$ are the average values over time of ln(TFR), respectively, ln(FLP) for country i in the sub-sample. In case of between-group estimation, one does not use time-series information to account for country effects. The third column in Table 1 shows the results of pooled least squares estimations with fixed country effects (approximated with a dummy variable for each country). The fourth column shows the results of generalized least squares estimations with random country effects. Furthermore, I allowed in case of fixed and random country effects estimation for the possibility of fixed time effects (approximated with a dummy variable for each time period). I included a dummy variable for each time period in case all time dummy variables were jointly significant according to a Chow test (in all of the tables below, this is indicated with either a "yes" or a "no"). Moreover, at the bottom of Table 1 and the following tables one can find test results of the null hypothesis of absence of country effects. In case of fixed effects estimation, this test is a Chow test with the null hypothesis of joint insignificance of the country dummy variables. In case of random effects estimation, this test is a Breusch-Pagan test with the null hypothesis that the variance of the random country effects equals zero. Fixed effects estimation is less efficient than random effects estimation due to a large loss in the degree of freedom. However, fixed effects estimation is more appropriate than random effects estimation if the country effects are correlated with the independent variable. The latter hypothesis can be tested with a Hausman test, which tests whether the residuals of pooled least squares estimations are correlated with the independent variable. It should be mentioned that for some countries, data, mostly for the FLP of certain years, were missing, i.e. the data set was an unbalanced panel. Where it was possible, I filled in missing values through interpolation.

Table 1 shows that, from 1960–1985, there was a negative and significant association between the TFR and the FLP, no matter which estimation method was applied. (In case of between-group estimation, the p-value is seven percent. This should be interpreted as significant, because the sample contains only twenty-one data points). The most important message of the table is that the absence of country effects can clearly be rejected (see the test results in the bottom row of the table).

Table 1. Explaining the TFR 1960–85, quinquennial data, unbalanced panel

Independent variables	Dependent variable: ln(TFR)		
	Between-group	Pooled LS with fixed country effects	Generalized LS with random country effects
Constant	0.56 (5.49)	0.69 (7.08)	0.73 (8.84)
Ln(FLP)	−0.25 (−1.94)	−0.38 (−3.56)	−0.35 (−4.16)
Fixed time effects included?		yes	yes
Number of observations	21	116	116
Specification tests		P-value	
H_0: Absence of country effects		0.00	0.00

Notes: t(z)-statistics are reported in parentheses below the coefficient estimates. Fixed time and country effects are not shown. R^2 within = 0.87, between = 0.16, overall = 0.65 (in case of fixed country effects estimation).

Table 2 contains the estimation results for the sub-sample 1985–2000, again with the TFR as the dependent variable and the FLP as the independent variable. Again, the table contains the results of between-group estimation, fixed country effects estimation and random country effects estimation (with fixed time effects, if significant). A glance at the second column in Table 2 reveals that with between-group estimation, the association between the TFR and the FLP is positive and significant. This result is consistent with the earlier findings of a positive cross-country association in post-1985 data. However, the next column in the table shows that with fixed country effect estimation, the association is negative and significant, while with random country effect estimation, the association is negative and insignificant. The bottom row of Table 2 shows results of a test Ho: Residuals not correlated with independent variable. This test is a Hausman test. The p-value of almost zero percent means that the test rejects the null hypothesis. This means that it suggests fixed country effects estimation. (In the following tables, results of a Hausman test are only shown and mentioned in the text, if the qualitative results differ between fixed and random country effects estimation). Moreover, the specification test at the bottom of the table above the Hausman tests shows that the absence of country effects can be rejected.

As a result, Table 2 shows that the association between the TFR and the FLP only changes its sign if one does not account for country effects. If one does accounts for country effects, however, the sign of the association remains negative. As is well known in econometrics literature, fixed country effects estimation is identical to within-group estimation. In turn, within-group estimation is pooled least squares regression of $[\ln(TFR_{i,t}) - \overline{\ln(TFR_i)}]$ on $[\ln(FLP_{i,t}) - \overline{\ln(FLP_i)}]$ (with, as already defined before, $\overline{\ln(TFR_i)}$, respectively, $\overline{\ln(FLP_i)}$ as the average values over time for country i in the sub-sample). This implies that in the case of fixed country effects estimation, the coefficient of ln(FLP) represents the time-series association between the TFR and the FLP within OECD countries (in contrast, random effects estimation contains

Table 2. Explaining the TFR 1985–2000, quinquennial data, unbalanced panel

Independent variables	Dependent variable: ln(TFR)		
	Between- group	Pooled LS with fixed country effects	Generalized LS with random country effects
Constant	0.66 (8.26)	0.33 (5.25)	0.47 (8.03)
Ln(FLP)	0.32 (2.31)	−0.28 (−2.40)	−0.03 (−0.32)
Fixed time effects included?		no	no
Number of observations	21	80	80
Specification tests		P-value	
H_0: Absence of country effects		0.00	0.00
H_0: Residuals not correlated with indep. variable		0.00	

Notes: t(z)-statistics are reported in parentheses below the coefficient estimates. Fixed time effects, and fixed and random country effects are not shown. R^2 within = 0.09, between = 0.22, overall = 0.10 (in case of fixed country effects estimation).

usually some cross-country information). Hence, applying fixed and random country effects estimation demonstrates that the *time-series association* between the TFR and the FLP does not change its sign, while applying between-group estimation demonstrates a reversal in the sign of the *cross-country association* between these two variables.[5]

In addition, comparing the results of Table 2 with those of Table 1 shows that the magnitude of the negative time-series association and the significance level was lower after 1985. This result is consistent with the view of Rindfuss et al. (2004), and Brewster and Rindfuss (2000) who argue that there has been a reduction in the incompatibility between childrearing and female employment due to institutional changes (including changing social norms). Appendix A contains formal statistical tests of whether there was a significant reduction in the negative time-series association between the TFR and the FLP. When the time-series association is assumed to be homogenous, then the tests give an ambiguous result. However, when the time-series association is allowed to be heterogeneous, just as in the next section, and the timing of the reduction is allowed to be different for different country groups, then the tests find unambiguous support for a reduction in the time-series association for countries that are neither Mediterranean nor Scandinavian countries.

4. Accounting for heterogeneity in the time-series association

The previous section has shown that, if one accounts for country effects, then the association between the TFR and the FLP does not change its sign. This is in contrast to the case without accounting for country effects where this association reverses its sign after about 1985. While this exercise demonstrated that the time-series association between the TFR and the FLP did not reverse its sign, it is unlikely to be able to explain the finding in the literature that the cross-country correlation changed its sign. Figure 2 in Sect. 2 illustrated a possible explanation for the reversal in the sign of the cross-country correlation between the TFR and the FLP. It was argued that this could be explained with the combination of country effects and, in addition, country-heterogeneity in the magnitude of the negative time-series association between fertility and female employment. Unfortunately, allowing the slope of ln(FLP) to be different for each country would lead to problems with difference-stationarity. This is so because ln(TFR) and ln(FLP) are not difference-stationary when pooled to a single sample, but are difference-stationary for each country alone. Most problematic, with difference-stationary data the standard errors are distorted, making inference on the significance of coefficients difficult. Therefore, I allow for heterogeneity in the slope of ln(FLP) for only three broad country groups: the Scandinavian countries (i.e., for Denmark, Finland, Norway and Sweden), the Mediterranean countries (i.e., for Greece, Italy, Portugal and Spain) and the remaining countries. It is well known that Scandinavian countries have a high TFR despite a high FLP, while the Mediterranean countries have a low TFR despite a low FLP. This section applies fixed and random country effects estimation with slope-heterogeneity for these three country groups (and with fixed time effects, if significant) to find out whether this heterogeneity can be confirmed in the data. Note that, while no longer discussed, specification

Table 3. Explaining the TFR with county-group heterogeneity, quinquennial data 1960–1985, unbalanced panel

Independent variables	Dependent variable: ln(TFR)	
	Pooled LS with fixed country effects	Generalized LS with random country effects
Constant	0.59 (6.16)	0.75 (8.74)
Dummy f. Scan. countries*	−0.04	−0.14
Ln(FLP)	(−0.28)	(−1.01)
Dummy f. Med. countries*	−0.97	−0.38
Ln(FLP)	(−3.52)	(−3.73)
Dummy f. other countries*	−0.49	−0.35
Ln(FLP)	(−4.34)	(−3.83)
Fixed time effects included?	yes	yes
Number of observations	116	116
Specification tests	P-value	
H_0: Absence of country effects	0.00	0.00

Notes: t(z)-statistics are reported in parentheses below the coefficient estimates. Fixed time effects, and fixed and random country effects are not shown. R^2 within = 0.89, between = 0.08, overall = 0.35 (in case of fixed country effects estimation).

tests – shown at the bottom of the following tables – confirm the presence of country effects.

Table 3 shows estimation results of quinquennial data for the sub-sample from 1960-85. The table confirm for fixed country effects estimation and for random country effects estimation differences in the coefficient of ln(FLP) between the three country groups.[6] As can be seen from the table the qualitative results are the same for fixed and random country effects estimation. As expected, the magnitude of the negative coefficient of ln(FLP) is the largest for Mediterranean countries and the lowest and even insignificant for Scandinavian countries.

Table 4 shows the corresponding estimation results for the sub-sample from 1985-2000. In this table, the qualitative estimation results differ somewhat between fixed and random country effects estimation. The table shows in the case of fixed country effects estimation a negative and significant coefficient of ln(FLP) for Mediterranean countries, while this coefficients is insignificant for Scandinavian countries and "the other countries" (i.e., countries that are neither Mediterranean nor Scandinavian countries). Regarding random country effects estimation, ln(FLP) is insignificant for all three country groups. However, a Hausman test (shown at the very bottom of the table) suggests fixed country effects estimation. Hence, upon taking into consideration of the results of the Hausman test, it follows that, after 1985, there was still a negative and significant time-series association between the TFR and the FLP for Mediteranean countries. In contrast, this time-series association was insignificant for Scandinavian countries and "the other countries" after 1985.

To conclude: Tables 3 and 4 show country-group-heterogeneity in the magnitude of the time-series association between the TFR and the FLP. The tables confirm the hypothesis of Fig. 2 that heterogeneity in this time-series association, along with the presence of country effects, might have caused a reversal of the cross-country association between the TFR and the FLP after about 1985.

Table 4. Explaining the TFR with county-group heterogeneity, quinquennial data 1985–2000, unbalanced panel

Independent variables	Dependent variable: ln(TFR)	
	Pooled LS with fixed country effects	Generalized LS with random country effects
Constant	0.39 (5.95)	0.46 (5.00)
Dummy f. Scan. countries*	0.98	−0.29
Ln(FLP)	(1.55)	(−0.94)
Dummy f. Med. countries*	−0.82	−0.09
Ln(FLP)	(−4.08)	(−0.71)
Dummy f. other countries*	−0.11	−0.07
Ln(FLP)	(−0.85)	(−0.53)
Fixed time effects included?	no	no
Number of observations	80	80
Specification tests	P-value	
H_0: Absence of country effects	0.00	0.00
H_0: Residuals not correlated with indep. variable	0.00	

Notes: t(z)-statistics are reported in parentheses below the coefficient estimates. Fixed time effects, and fixed and random country effects are not shown. R^2 within = 0.26, between = 0.38, overall = 0.26 (in case of fixed country effects estimation).

In addition, comparing for "the other countries" the results of Table 4 with those of Table 3 shows a reduction in the magnitude and the significance level of the time-series association for these countries. As mentioned before, Appendix A contains formal statistical tests of whether the negative time-series association fell significantly. The tests find for "the other countries" unambiguous support for a reduction in the time-series association, when for different country groups the timing of the reduction is allowed to differ. It should be kept in mind, that in reality there was probably a gradual reduction in the time-series association rather than a single break. However, the more important point is that the time-series association did significantly fall.

5. Conclusion

Recent research in e.g. Ahn and Mira (2002), and Rindfuss et al. (2004) found that the cross-country correlation between the TFR and the FLP in OECD countries, which had been negative until about 1985, had changed to a positive value since then. Rindfuss et al. (2004), and Brewster and Rindfuss (2000), point to changes in the institutional context, such as changing government policies, changing attitudes toward working mothers, and an increased availability of child-care. All are factors that reduced incompatibility between childrearing and female employment. However, Engelhardt et al. (2004) found in time-series data of six OECD countries that the negative time-series association between the TFR and the FLP became weaker and less significant over time for all non-Mediterranean countries in their study, but it did not change its sign for any country.

This paper applied econometric methods that account for country effects in pooled cross-country and time-series data of OECD countries. Data from

1960–2000 were divided into the sub-samples 1960–1985 and 1985–2000 in order to discover whether or not the association between the TFR and the FLP had changed its sign after 1985. In using this framework, the study has shown that the time-series association does not demonstrate a change in sign. However, the study shows that, for countries that are neither Mediterranean nor Scandinavian countries, the magnitude and the significance level of the time-series association were lower after 1985 than before. The finding of a falling magnitude and significance level of this association is consistent with the theoretical argument in Rindfuss, Benjamin and Morgan, and Brewster and Rindfuss of a falling incompatibility between childrearing and female employment. In addition, the paper finds heterogeneity in the magnitude of the negative time-series association between the TFR and the FLP for three broad country groups. Most importantly, it shows that the magnitude of this negative time-series association was the largest for Mediterranean countries and the smallest for Scandinavian countries. The paper shows that the presence of country effects and heterogeneity in the magnitude of the negative time-series association between fertility and female employment together very likely explain the finding of a reversal in the sign of the cross-country association between the TFR and the FLP.

Appendix A: Tests for a break in the time-series association between the TFR and the FLP

This section shows results of regressions of the TFR on the FLP and the other way around. In order to check whether the negative time-series association between the TFR and the FLP became statistically significantly weaker over time, Tables A1–A5 contain formal tests for a break in the slope of the FLP, respectively, the TFR. Tables A1 and A2 contain results of tests for a break in the slope of ln(FLP), respectively, ln(TFR) in 1985. These tests assume a homogeneous slope of ln(FLP), respectively, ln(TFR), just as in Sect. 3 for ln(FLP). Tables A3 and A4 contain results of tests that allow for a heterogeneous slope of ln(FLP), respectively, ln(TFR) for three country groups, just as in Sect. 4 for ln(FLP), and which test for a common break of this slope in 1985. In contrast to this, Table A5 contains results of tests where the dates of the breaks are allowed to be different for each country group and the dates of the breaks in the slopes are endogenously chosen (according to a procedure explained below).

Tables A1–A5 contain results of fixed and random country effect estimation with quinquennial data from 1960–2000 (and fixed time effects included, if significant). As the time-dimension of the sample of Tables A1–A5 is larger than in Tables 1–4, the residuals are first order autoregressive in case of estimation with these data. This can be seen from the Baltagi-Wu LBI-statistics (which are shown in the notes below Tables A1–A5). The Baltagi-Wu LBI-statistic is the equivalent of the Durbin-Watson statistic and is the relevant statistic for a test of serial correlation in the case of an unbalanced panel (because the Durbin-Watson-statistic is not appropriate in case of an unbalanced panel). A value of the Baltagi-Wu LBI-statistic far below 2 indicates that correction for serial correlation is clearly necessary (exact critical values are not available in the literature). Because of Baltagi-Wu LBI-statistics far below 2 in all cases of Table A1-A5, estimation is in all cases

Table A1. Testing for presence of break in slope with the TFR as dependent variable, 1960–2000, quinquennial data, unbalanced panel

Independent variables	Dependent variable: ln(TFR)	
	Pooled LS with fixed country effects	Generalized LS with random country effects
Constant	0.17 (5.89)	0.71 (8.69)
Ln(FLP)	−0.41 (−2.78)	−0.35 (−4.13)
Dummy for 1985–2000* Ln(FLP)	0.36 (4.30)	0.39 (4.85)
Fixed time effects included?	yes	yes
Number of observations	154	175

Notes: t(z)-statistics are reported in parentheses below the coefficient estimates. Fixed time effects, and fixed and random country effects are not shown. Estimation with first-order autoregressive residuals according to the method of Baltagi and Wu (1999). In case without correction for autocorrelation the Baltagi-Wu LBI-statistic was 1.26 for fixed country effects estimation and 1.21 for random country effects estimation. R^2 within = 0.53, between = 0.05, overall = 0.59 (in case of fixed country effects estimation).

Table A2. Testing for presence of break in slope with the FLP as dependent variable, 1960-2000, quinquennial data, unbalanced panel

Independent variables	Dependent variable: ln(FLP)	
	Pooled LS with fixed country effects	Generalized LS with random country effects
Constant	−0.07 (−8.22)	−0.70 (−9.28)
Ln(TFR)	−0.07 (−1.27)	−0.18 (−2.98)
Dummy for 1985–2000* Ln(TFR)	−0.06 (−0.87)	0.08 (0.99)
Fixed time effects included?	yes	yes
Number of observations	154	175

Notes: t(z)-statistics are reported in parentheses below the coefficient estimates. Fixed time effects, and fixed and random country effects are not shown. Estimation with first-order autoregressive residuals according to the method of Baltagi and Wu (1999). In case without correction for autocorrelation the Baltagi-Wu LBI-statistic was 1.11 for fixed country effects estimation and 0.96 for random country effects estimation. R^2 within = 0.41, between = 0.00, overall = 0.17 (in case of fixed country effects estimation).

applied with first order autoregressive residuals, according to the method of Baltagi and Wu (1999).

Regarding the case with a homogenous slope of ln(FLP), respectively, ln(TFR), in Table A1, the TFR is the dependent variable and in Table A2, the FLP is the dependent variable. In addition, Table A1 includes the interaction variable “Dummy for 1985–2000*ln(FLP)”. The variable “Dummy for 1985–2000” has the value one for 1985, 1990, 1995 and 2000, and zero otherwise and “*” denotes multiplied. Similarly, Table A2 includes the interaction variable “Dummy for 1985–2000*ln(TFR)”. Table A1 demonstrates a statistically significant reduction in the time-series association between the TFR and the FLP in the case of the TFR as the dependent variable. In this table, the interaction variable “Dummy for 1985–2000*ln(FLP)” is positive and significant according to the t-statistic for fixed country effects estimation and

Table A3. Testing for presence of breaks in slope with the TFR as dependent variable, 1960–2000, quinquennial data, unbalanced panel

Independent variables	Dependent variable: ln(TFR)	
	Pooled LS with fixed country effects	Generalized LS with random country effects
Constant	0.24 (6.57)	0.76 (8.93)
Dummy f. Scan. countries* ln(FLP)	0.15 (0.45)	−0.22 (−1.45)
Dummy f. Med. countries* ln(FLP)	−0.92 (−4.72)	−0.36 (−3.92)
Dummy f. other countries* ln(FLP)	−0.13 (−0.81)	−0.31 (−3.36)
Dummy for 1985–2000*dummy f. Scan. countries*ln(FLP)	0.31 (1.11)	0.10 (0.41)
Dummy for 1985–2000*dummy f. Med. countries*ln(FLP)	0.34 (3.33)	0.41 (4.22)
Dummy for 1985–2000*dummy f. other countries*ln(FLP)	0.28 (2.33)	0.22 (1.88)
Fixed time effects included?	yes	yes
Number of observations	154	175

Notes: t(z)-statistics are reported in parentheses below the coefficient estimates. Fixed time effects, and fixed and random country effects are not shown. Estimation with first-order autoregressive residuals according to the method of Baltagi and Wu (1999). In case without correction for autocorrelation the Baltagi-Wu LBI-statistic was 1.36 for fixed country effects estimation and 1.32 for random country effects estimation. R^2 within $= 0.66$, between $= 0.12$, overall $= 0.06$ (in case of fixed country effects estimation).

for random country effects estimation, as well. In contrast, Table A2 reveals a statistically insignificant reduction in this time-series association in the case of the FLP as the dependent variable. In this table, the interaction variable "Dummy for 1985–2000*ln(TFR)" is insignificant according to the t-statistic, for both, fixed and random country effects estimation. Hence, a formal test of whether the time-series association between the TFR and the FLP fell significantly gives an ambiguous result (depending on which variable is the dependent variable).

Regarding the case with a heterogeneous slope of ln(FLP), respectively, ln(TFR), and an exogenously chosen date of the possible break in 1985, the TFR is the dependent variable in Table A3 and the FLP is the dependent variable in Table A4. In addition, Table A3 includes the interaction variable "Dummy for 1985–2000*dummy for country group z*ln(FLP)" for each country group z, where the country groups are: the Scandinavian, the Mediterranean or "the other" countries. Similarly, Table A4 includes the interaction variable "Dummy for 1985–2000*dummy for country group z*ln(TFR)" for each of the three country groups z. Table A3 shows similar results for fixed and random country effects estimation. No matter which estimation method was applied, there is always a positive and statistically significant coefficient of the interaction variables "Dummy for 1985–2000*dummy f. Med. countries*ln(FLP)" and of "Dummy for 1985–2000*dummy f. other countries*ln(FLP)", while the interaction variable "Dummy for 1985–2000*dummy f. Scan. countries*ln(FLP)" is insignificant. This implies for Mediterranean countries and "the other countries" a

Table A4. Testing for presence of breaks in slope with the FLP as dependent variable, 1960-2000, quinquennial data, unbalanced panel

Independent variables	Dependent variable: ln(FLP)	
	Pooled LS with fixed country effects	Generalized LS with random country effects
Constant	−0.08 (−8.32)	−0.69 (−9.06)
Dummy f. Scan. countries* ln(TFR)	−0.09 (−0.80)	−0.06 (−0.65)
Dummy f. Med. countries* ln(TFR)	−0.19 (−1.70)	−0.30 (−3.43)
Dummy f. other countries* ln(TFR)	0.01 (0.12)	−0.19 (−2.78)
Dummy for 1985–2000*dummy f. Scan. countries*ln(TFR)	−0.06 (−0.65)	0.15 (1.44)
Dummy for 1985–2000*dummy f. Med. countries* ln(TFR)	−0.02 (−0.28)	0.03 (0.29)
Dummy for 1985–2000*dummy f. other. countries*ln(TFR)	−0.06 (−0.88)	0.09 (1.05)
Fixed time effects included?	yes	yes
Number of observations	154	175

Notes: t(z)-statistics are reported in parentheses below the coefficient estimates. Fixed time effects, and fixed and random country effects are not shown.
Estimation with first-order autoregressive residuals according to the method of Baltagi and Wu (1999). In case without correction for autocorrelation the Baltagi-Wu LBI-statistic was 1.14 for fixed country effects estimation and 0.99 for random country effects estimation. R^2 within = 0.43, between = 0.00, overall = 0.09 (in case of fixed country effects estimation).

Table A5. Testing for presence of *endogenous* breaks in slope with the FLP as dependent variable, 1960–2000, quinquennial data, unbalanced panel

Independent variables	Dependent variable: ln(FLP)	
	Pooled LS with fixed country effects	Generalized LS with random country effects
Constant	−0.09 (−12.49)	−0.73 (−10.71)
Dummy f. Scan. countries* ln(TFR)	−0.11 (−1.21)	−0.02 (−0.20)
Dummy f. Med. countries* ln(TFR)	−0.24 (−2.41)	−0.25 (−3.29)
Dummy f. other countries* ln(TFR)	−0.01 (−0.20)	−0.15 (−2.48)
Dummy for 1980–2000*dummy f. Med. countries*ln(TFR)	−0.06 (−1.36)	−0.07 (−1.31)
Dummy for 1995–2000*dummy f. other. countries*ln(TFR)	0.14 (3.00)	0.14 (2.57)
Fixed time effects included?	yes	yes
Number of observations	154	175

Notes: t(z)-statistics are reported in parentheses below the coefficient estimates. Fixed time effects, and fixed and random country effects are not shown. Estimation with first-order autoregressive residuals according to the method of Baltagi and Wu (1999). In case without correction for autocorrelation the Baltagi-Wu LBI-statistic was 1.13 for fixed country effects estimation and 0.96 for random country effects estimation. R^2 within = 0.46, between = 0.01, overall = 0.17 (in case of fixed country effects estimation).

statistically significant reduction in the time-series association for the case with the TFR as the dependent variable. Also Table A4 shows similar results for fixed and random country effects estimation. However, Table A4 shows an insignificant coefficient of the interactions variable "Dummy for 1985–2000*dummy for country group z*ln(FLP)" for all three country groups z. Hence, in the case of the FLP as the dependent variable, there is for all three country groups no statistically significant reduction in the time-series association between the TFR and the FLP. As a consequence, Tables A3 and A4 together imply for the Mediterranean and "the other" countries an ambiguous result. (For Scandinavian countries there is clearly no statistical reduction in the time series-association between these two variables).

Finally, the tests of Table A5 repeated the tests of Table A4. However, in the tests of Table A5, the dates of the breaks in the slopes are allowed to be different for different country groups. In addition, the dates of these breaks were chosen endogenously. Regressions of the type of Table A4 were repeated for various values of the date t_B for the interaction variables "Dummy for time period $t_B - 2000$*dummy for countries group z*ln(TFR)" for all three country groups z (where $t_B - 2000$ means from year t_B to year 2000). For each country group, the value of the date t_B for which the absolute value of the t-statistic of the interaction variable is maximized was chosen as the "optimal" break date (i.e., the date of the break with the best fit in the data). Table A5 shows similar results for fixed and random country effects estimation. In case of both estimation methods, the "optimal" dates of the break are 1980 for the Mediterranean countries and 1995 for "the other countries" (in the tests of Table A5 no break is assumed for the Scandinavian countries, because Table A4 demonstrated that for Scandinavian countries ln(TFR) is insignificant in the entire time period 1960–2000). Further, in case of both estimation methods, the coefficient of the interaction variable is insignificant for Mediterranean countries and positive and significant for "the other countries". To conclude: If the date of the break in the slope of ln(TFR) is allowed to vary between the three country groups and is endogenously chosen for each country group, then formal tests show unambiguous support for a reduction in the time-series association between the TFR and the FLP for "the other countries".

Appendix B: Estimation results with FLP of women aged 25 to 44

A drawback of Tables 1–4 is that the FLP contains all women aged 15 to 64, i.e., women above the age of 44 who do not bear children and often no longer rear children. Hence, by using the FLP of ages 15 to 64, one can only approximate a measure for labor force participation for women of childrearing age. Table B1 shows the results of estimation, with the FLP of women aged 15 to 44 instead of women aged 15 to 64. As such data do not exist prior to 1970, these tables show only estimation results with quinquennial data for our second sub-sample 1985–1995 (data for 2000 are not yet available).

Just as the tables in Sect. 3, Table B1 shows results of between-group estimation, fixed country effects estimation, and random country effect estimation, with the TFR as the dependent variable and a homogeneous slope of ln(FLP) (and again results with the FLP as the dependent variable are very similar and therefore are not shown). The second column in the table shows that, with between-group estimation (and therefore without accounting for country effects), the association between fertility and female employment is again positive (although not very significantly). However, specification tests (shown in the row above the bottom row) reveal that, also with these data, the absence of country effects can be rejected. Thus, one needs to account for country effects. The third and fourth columns show the results of fixed and random country effects estimation. The table shows that the time-series association is negative and significant for fixed country effects estimation and is negative and insignificant for random country effects estimation. A Hausman test (results are shown at the very bottom) suggests the use of fixed country effects estimation. If the reader compares the results of Table B1 with those in Table 2, then he/she will realize that with the FLP of women aged 25 to 44, there is more support for a significant negative time-series association between the TFR and the FLP after 1985 than with the FLP of women aged 15 to 64. Hence, with the FLP of women aged 25 to 44, the hypothesis of this paper is even reinforced.

Finally, Tables B2 shows results where the tests of Tables 4, i.e. tests with a heterogeneous slope of ln(FLP) are repeated with data of the FLP of women aged 25 to 44, instead of the FLP of women aged 15 to 64. Again, only results with the TFR as the dependent variable are shown, as the results with the FLP as the dependent variable were very similar. Regarding fixed country effects estimation, the table shows a negative and significant coefficient of the FLP of women aged 25–44 for Mediterranean countries and an insignificant coefficient for Scandinavian countries and "the other countries". Regarding random country effects estimation, this coefficient is only negative and significant for "the other countries". Results from a Hausman test are shown at the bottom of the table. They suggest the use of fixed country effects estimation. Hence, upon use of the Hausman test result, one can conclude, that only for Mediterranean countries, there was a negative and significant time-

Table B1. Explaining the TFR, 1985–1995, quinquennial data, unbalanced panel

Independent variables	Dependent variable: ln(TFR)		
	Between-group	Pooled LS with fixed country effects	Generalized LS with random country effects
Constant	0.60 (8.48)	0.34 (6.16)	0.45 (8.25)
Ln female employment of women 25—44	0.23 (1.38)	−0.42 (−3.05)	−0.15 (−1.34)
Fixed time effects included?		no	no
Number of observations	18	53	53
Specification tests		P-value	
H_0: Absence of country effects		0.00	0.00
H_0: Residuals not correlated with indep. Variable		0.00	

Notes: t(z)-statistics are reported in parentheses below the coefficient estimates. Fixed time effects, and fixed and random country effects are not shown. R^2 within = 0.21, between = 0.11, overall = 0.02 (in case of fixed country effects estimation).

Table B2. Explaining the TFR with county-group heterogeneity, quinquennial data 1985–1995, unbalanced panel

Independent variables	Dependent variable: ln(TFR)		
	Pooled LS with fixed country effects		Generalized LS with random country effects
Constant	0.37 (6.48)		0.41 (6.94)
Dummy f. Scan. countries* ln FLP of women age 25–44	0.24 (0.29)		−0.76 (−1.77)
Dummy f. Med. countries* ln FLP of women age 25–44	−0.84 (-3.94)		−0.02 (−0.14)
Dummy f. other countries* ln FLP of women age 25–44	−0.19 (-1.15)		−0.28 (−2.15)
Fixed time effects included?	no		no
Number of observations	53		53
Specification tests		P-value	
H_0: Absence of country effects	0.00		0.02
H_0: Residuals not correlated with indep. variable		0.00	

Notes: t(z)-statistics are reported in parentheses below the coefficient estimates. Fixed time effects, and fixed and random country effects are not shown. R^2 within = 0.35, between = 0.47, overall = 0.24 (in case of fixed country effects estimation).

series association after 1985 between the TFR and the FLP. A comparison of the results in Tables B2 with those of Tables 4 shows again very similar results with the FLP of women aged 25–44 and with the FLP of women aged 15–64.

Appendix C: Panel data unit root tests

This Appendix shows the results of panel data unit root tests for all of the variables in the paper. As is explained in the text, if the series contained a unit root, then standard inference of the estimation results were not possible.

Recently, Harris and Tzavalis (1999) derived critical values for Dickey-Fuller (DF) tests of pooled series with a large cross-section dimension and a small time-series dimension. They consider the cases of: (i) a homogenous panel, (ii) a panel with fixed effects for the mean (i.e., inclusion of country dummy variables in the unit root test equations) and without a deterministic trend, and (iii) a panel with fixed effects in the mean and individual deterministic trends (where the latter has the value one in 1960, two in 1965 and so forth). Applying Chow tests, I could not reject the absence of fixed effects in the mean for any variable in the paper. Hence, country dummy variables were included in all unit root tests. Further, individual deterministic trends were included, if they were jointly significant according to a Chow test.

Harris and Tzavalis show that the limiting distribution of the test statistic is normal. This means that one can apply standard DF tests to the pooled series and can use the standard t-statistic-criterium for inferring whether or not the series of consideration contains a unit root. (The exact critical values are shown in various tables in Harris and Tzavalis). In DF tests, one applies OLS regressions of the first difference of a series on its lagged level. If the lagged level is negative and significant, the presence of a unit root in the level is rejected. The test statistic of the lagged level is often referred to as a DF statistic.

Table C1. Panel data unit root test with quinquennial data

Series, sample	DF	Trend included?
Ln(TFR), 1960–85	−7.12*	yes
Ln(FLP, 15–64), 1960–85	−6.38*	yes
Ln(TFR), 1985–2000	−8.41*	no
Ln(FLP, 15–64), 1985–2000	−11.26*	yes
Ln(TFR), 1960–2000	−5.26*	no
Ln(FLP, 15–64), 1960–2000	−5.52*	yes
Ln(TFR), 1985–1995	−5.43*	no
Ln (FLP, 25–44), 1985–1995	−5.54*	no

Notes: A "*" denotes significant at 5% level.
FLP, 15–64 = FLP of women of age 15–64, FLP, 25–44 = FLP of women of age 25–44.

Table C1 shows the DF test statistics for all of the variables in the paper (in addition, Table C1 includes information whether individual, deterministic trends were included in the unit root test of a particular variable). A glance at Table C1 reveals that the lagged level of all the series in this study is negative and significant. Hence, none of these variables contains a unit root.

Appendix D: List of countries and data

Countries in Fig. 1 and all tables (for some time periods, data were missing and in Table B1–B4 Austria, Switzerland, and Western-Germany were not included due to lack of enough time-series data for age-specific FLP's):

Australia	Greece	Portugal
Austria	Ireland	Spain
Belgium	Italy	Sweden
Canada	Japan	Switzerland
Denmark	Luxembourg	United Kingdom
Finland	Netherlands	United States
France	Norway	Western-Germany

TFR: Total fertility rate

Definition: Sum of age specific fertility rates.

Data sources: For all European countries the source is "Recent demographic developments in Europe 2001", Council of Europe. For all non-European countries the sources are (for 1960–1985) "UN Demographic yearbook 1948–1997", CD-ROM (for Australia also for 1996) and (for 1996–1999) for the USA and Japan "New Cronos 2001" (Eurostat Database). Further, for all non-European countries the sources for 2000 are: Australian Bureau of Statistics for Australia, National Institute of Population and National Security Research Japan "Latest Demographic Statistics" for Japan, and US Census Bureau for the USA.

FLP: Female labor force participation rate of women ages 15–64

Definition: Female labor force of women of all ages, including unemployed women of that age, divided by female population of age 15 to 64.

Data source: Comparative welfare states and OECD Labour Force Statistics (1997, 1998 and 2001) for all countries, except Western-Germany after 1989, where the source is: German Federal Statistical Office, micro-census.

Note: In case of Norway, in OECD Labor Force Statistics, data after 1970 were from the labor force survey and before 1970 from the population and household census. As the data before 1970 differ very much from the data after 1970, I used for Norway only data since 1970. Further, I did not use any data of New Zealand. The reason for this is the fact that, in OECD Labor Force Statistics, data after 1985 were collected from Statistics New Zealand and before 1985 they were collected from the Department of Labor in New Zealand and the data before 1985 differed very much from those after 1985. I omitted New Zealand completely (instead of using data since 1985), because in the paper I splitted the sample into the sub-samples 1960–1985 and 1985–2000 and both samples should contain the same countries.

FLP, 25–44: Female labor force participation rate of women ages 25–44

Definition: Female labor force of women of age 25 to 44 including unemployed women of that age divided by female population of age 25 to 44.

Data sources: FLP of women of age 25–34 and 35–44 from OECD Labour Force Statistics (2001). The FLP of women of age 25–44 was calculated by weighting the two aforementioned age-specific FLP's by the share of female population of that age in the female population of age 25–44. The data source of age-specific female population was U.N. Demographic Yearbook 1948–1997, CD-ROM.

Endnotes

[1] A positive association between fertility and female employment in macro-level data since the 1980s was already suggested earlier in Bernhardt (1993), Pinelli (1995) and Rindfuss and Brewster (1996).

[2] The countries in their study were France, Italy, Sweden, the United Kingdom, the United States, and Western-Germany.

[3] I am grateful to a referee for suggesting that I show a figure such as Fig. 2 to illustrate my hypothesis.

[4] Contrary to Fig. 2, in the more rigorous exercises in Sects. 3 and 4 country effects had also on the TFR a positive effect in high fertility countries and a negative effect in low fertility countries.

[5] I am grateful to an anonymous referee for suggesting this interpretation of the evidence of Table 2.

[6] In Table 3 and the following Table 4 the dummies are "group" slope coefficients, and in addition to them the regressions contain fixed or random country effects.

References

Ahn N, Mira P (2002) A Note on the Changing Relationship between Fertility and Female Employment Rates in Developed Countries. *Journal of Population Economics* 15(4):667–682

Apps A, Rees R (2001) Fertility, Female Labor Supply and Public Policy. Discussion Paper No. 409, Institute of the Study of Labor (IZA), Bonn, Germany

Baltagi BH, Wu P (1999) Unequally Spaced Panel Data Regression with AR(1) Disturbances. *Econometric Theory* 15(6):814–823

Benjamin K (2001) Men, Women and Low Fertility: Analysis across Time and Country. Paper presented at the 2001 Annual Meeting of the Population Association of America, Washington, DC

Bernhardt, EM (1993) Fertility and Employment. *European Sociological Review* 9(1):25–42

Brewster KL, Rindfuss RR (2000) Fertility and Women's Employment in Industrialized Nations. *Annual Review of Sociology* 26:271–296

Cramer JC (1980) Fertility and Female Employment: Problems of Causal Directions. *American Sociological Review* 45(2):167–190

Engelhardt H, Kögel T, Prskawetz A (2004) Fertility and female employment reconsidered: a macro-level time series analysis for developed countries, 1960–2000. *Population Studies* (forthcoming)

Esping-Andersen G (1999) Social Foundations of Postindustrial Economies. Oxford University Press, Oxford

Galor O, Weil DN (1996) The Gender Gap, Fertility, and Growth. *American Economic Review* 86(3):374–387

Granger CWJ, Newbold P (1974) Spurious Regression in Econometrics. *Journal of Econometrics* 2(2):111–120

Harris RDF, Tzavalis E (1999) Inference for Unit Roots in Dynamic Panels where the Time Dimension is fixed. *Journal of Econometrics* 91(2):201–226

Lehrer E, Nerlove M (1986) Female Labor Force Behavior and Fertility in the United States. *Annual Review of Sociology* 12:181–204

Martinez DF, Iza A (2004) Skill Premium Effects on Fertility and Female Labor Supply. *Journal of Population Economics* 17:1–16 (this issue)

Pinelli, A (1995) Women's Condition, Low Fertility, and Emerging Union Patterns in Europe. In: K Oppenheim, KO Mason, A-M Jensen (eds). *Gender and Fertility Change in Industrialized Countries*. Clarendon Press, Oxford, 82–101

Rindfuss RR, Guzzo KB, Morgan SP (2004) The Changing Institutional Context of Low Fertility. *Population Research and Policy Review* (forthcoming)

Rindfuss RR, Brewster KL (1996) Childrearing and Fertility. *Population and Development Review* (Supplement) 22:258–289