Open access • Journal Article • DOI:10.1007/S10888-013-9260-9

Inequality and the time structure of earnings: evidence from Germany

Carsten Schröder, Yolanda Golan, Shlomo Yitzhaki

Institutions: Kiel Institute for the World Economy, Central Bureau of Statistics, Hebrew University of Jerusalem

Published on: 01 Sep 2014 - Journal of Economic Inequality (Springer US)

Topics: Earnings, Accounting period, Economic inequality and Time horizon

Schröder, Carsten; Golan, Yolanda; Yitzhaki, Shlomo

Article — Accepted Manuscript (Postprint)

Inequality and the time structure of earnings: Evidence from Germany

Journal of Economic Inequality

Provided in Cooperation with:
German Institute for Economic Research (DIW Berlin)

Suggested Citation: Schröder, Carsten; Golan, Yolanda; Yitzhaki, Shlomo (2014) : Inequality and the time structure of earnings: Evidence from Germany, Journal of Economic Inequality, ISSN 1569-1721, Springer, Heidelberg, Vol. 12, Iss. 3, pp. 349-409, https://doi.org/10.1007/s10888-013-9260-9

The final publication is available at www.springerlink.com:
http://dx.doi.org/10.1007/s10888-013-9260-9

Inequality and the time structure of earnings: Evidence from Germany

Carsten Schröder, University of Kiel and Kiel Institute for the World Economy, Germany, carsten.schroeder@economics.uni-kiel.de (for correspondence and/or offprints)

Yolanda Golan, The Central Bureau of Statistics, Israel, yolandag@cbs.gov.il

Shlomo Yitzhaki, Department of Economics, Hebrew University Jerusalem, Israel, shlomo.yitzhaki@huji.ac.il

Abstract. This paper studies the relationships between annual and subannual inequality and mobility during the course of the year. We apply an exact decomposition framework as outlined in Wodon and Yitzhaki (2003), and in Yitzhaki and Wodon (2004). Earnings records of pension insurants in Germany serve as the database. The long time horizon of our database allows us to investigate the stability and robustness of the parameters of the decomposition over time. Specifically, we show that the mobility component of the decomposition, as measured by Gini correlation coefficients, changes over the observation period. This makes it difficult to predict the impact of the income accounting period on inequality in a more general context. Thus, it is of paramount importance to use income data from a uniform accounting period in distributional analyses.

Keywords. Accounting period; time structure of earnings; Gini decomposition; inequality; mobility; re-ranking

1 Introduction

There is an increasing interest in issues of distributive justice, fairness, and equity. The Canberra Group (2001) has provided detailed guidelines for compiling micro databases that allow valid comparative distributional analyses. Institutions like the Luxembourg Income Study have compiled micro databases from numerous countries, and invested substantial efforts to ensure cross-country comparability of the information content of the databases. As a result, various national and international micro databases are available to study distribution issues.

Despite such efforts, a potential source of heterogeneity remains: the income accounting period (IAP).[1] The IAP is the measurement period of the flow variable income, and it can be defined as a calendar or fiscal period. Typically, it is a year, a quarter, or a month. If IAPs differ, results from comparative distributional analysis are likely to be biased. As an example, inequality and the IAP will usually be negatively related, because an extension of the IAP smoothes subannual income fluctuations. Accordingly, measuring a higher level of inequality may be an artifact resulting from different IAPs rather than from true comparison. Ideally, therefore, distributional analysis should rest on incomes covering a uniform accounting period.

Should IAPs differ, valid comparative distributional analyses still would be possible if a general rule existed for how IAPs impact distribution indices. However, such a general rule, like "halving the IAP increases measured inequality by five percent," does not exist. Shorrocks (1978a) has shown that inequality and the IAP are negatively related under general conditions. The quantitative strength of the relationship, however, is not determined. For the impact of the IAP

[1] Even highly standardized databases like the Luxembourg Income Study (LIS) rely on data from accounting periods of different lengths. For the French LIS database, the IAP is the "twelve months preceding the interview," for Australia "the financial year preceding the date of interview," for Germany and the US the "calendar year" (see original survey information provided at http://www.lisdatacenter.org/ for details). The inter-temporal comparability of some national databases is limited by IAP adjustments over time. As an example, a peculiarity of the income and expenditure surveys for Israel and Germany is that the accounting period has been shortened from an annual to a quarterly time span. In Israel, the accounting period now is three months, but the survey is conducted over the year. Each time the question refers to the income earned in the last three months, e.g., respondents surveyed in January are asked about October to December of the previous year. In the German survey, conducted in five-year intervals, households provide earnings and expenses for various kinds of goods and services in household diaries. Until 1993, information from these diaries was collected over a period of twelve months. Since 1998, the survey has been conducted over a full year but each household fills in its diary for a single, random quarter in order to lessen the effort involved for the participant, to improve the data quality, and to reduce drop-out rates (see Kühnen, 2001).

on poverty or mobility, the picture is even less clear. Here, the direction of the relationship is usually undetermined (see Böheim and Jenkins, 2006; Ravallion, 1988; Chesher and Schluter, 2002).

Because theory provides little guidance for sorting out the effect of the IAP on distributional measures, empirical studies of the IAP-distribution nexus are of paramount importance. However, the data requirement for a systematic empirical assessment is high, explaining the scarcity of empirical evidence. Ideally, an empirical analysis of the IAP-distribution nexus requires micro data on incomes for a short accounting period (e.g., a month) available for a long time horizon (e.g., a decade), allowing monthly distributions to be derived. Further, by adding up monthly incomes over time (e.g., a year), the associated longer-term (annual) distributions could be derived as well. Together, these distributions would allow for a systematic analysis of the relationship between the accounting period and distributional measures.

Empirical evidence on the IAP-distribution nexus is provided in Shorrocks (1981) and Ruggles (1990) for the United States, Morris and Preston (1986), Nolan (1987) and Böheim and Jenkins (2000, 2006) for the United Kingdom, Gibson et al. (2001) for China, Cantó et al. (2006) for Spain, Finkel et al. (2006) for Israel, and Detlefsen (2012) and Schröder (2012) for Germany. Effectively, however, the empirical evidence is smaller than the mere number of studies suggests. This is because the IAP-distribution nexus is sometimes investigated as a byproduct, or because data restrictions prohibit a rigorous empirical analysis (for details see Böheim and Jenkins, 2006; Cantó et al., 2006; Schröder, 2012). Further, none of the studies, except for Finkel et al. (2006), uses a framework that enables a systematic analysis of the IAP-distribution nexus.

For these reasons, Finkel et al. (2006), Wodon and Yitzhaki (2003), and Yitzhaki and Wodon (2004) argue that further research on the IAP-distribution nexus is required. This paper addresses that challenge. Employment records of German pension insurants, providing monthly information on earnings, serve as our database.

This paper contributes by providing systematic empirical evidence on the IAP-distribution nexus over a long time horizon. We apply an exact decomposition framework, as suggested in Wodon and Yitzhaki, 2003, and in Yitzhaki and Wodon, 2004, that decomposes the square of the Gini index from the annual distribution in three component: (a) a series of Gini indices from the monthly distributions; (b) a series of Gini correlation coefficients[2] of the monthly distributions amongst themselves;[3] and (c) differences between the two Gini correlations defined between each monthly and the yearly distribution. These differences describe whether the yearly distribution and the monthly distributions belong to the same family of distributions. The long time horizon of our database allows us to investigate the inter-temporal stability and robustness of the parameters (in the same country).

Our results show that inequality indices from the subannual distributions are significantly larger than indices from the annual distributions. Specifically, the Gini coefficient of the monthly distributions is 2.2 percent higher, on average, than the coefficient of the annual distribution. Despite such differences in absolute levels, Gini indices from both annual and monthly earnings distributions follow the same inter-temporal pattern, indicating a significant rise in earnings inequality since German reunification. Finally, Gini correlations of the monthly distributions vary over the observation period. This time-variance makes it difficult to predict the quantitative impact of the IAP on inequality in a more general context. Thus, it would appear that it is necessary to have a uniform IAP to derive valid conclusions from any comparative distributional analysis.

The structure of the paper is as follows. Section 2 explains the Gini decomposition framework and outlines the statistical procedures. Section 3 describes the database and its preparation. Section 4 provides our estimates on the IAP-inequality nexus, and Section 5 concludes.

2 Methods

2.1 The Gini decomposition framework

Our empirical analysis builds on the Gini decomposition framework introduced in Wodon and Yitzhaki (2003) and in Yitzhaki and Wodon (2004). This framework enables a systematic investigation of the relationships between the Gini index derived from the annual distribution,

[2] The Gini correlation coefficients, introduced in Schechtman and Yitzhaki (1987), are based on the covariance between one income variable and the cumulative distribution of another income variable. Gini correlation coefficients, like Pearson's or Spearman's correlations, describe the dependence between two variables.

[3] The Gini decomposition framework, in the spirit of the works of Atkinson (1983), Shorrocks (1978b and c), King (1983), Atkinson and Bourguignon (1992), or Dardanoni (1993), integrates mobility as an additional dynamic dimension describing distributional transition processes.

and the Gini indices and the Gini correlations derived from the subannual monthly distributions. A particular advantage of the Gini decomposition framework is that "mobility is not defined as an independent concept, and therefore, there is no need to derive a separate axiomatic justification for it" (Yitzhaki and Wodon, 2004, p. 181).

Gini indices and Gini correlation coefficients constitute the basic ingredients of the Gini decomposition framework. Let $Z_1,...,Z_M$ denote the earnings distribution in M periods (here: January to December). Further, let $Y_m = Z_m / \mu_{Z_m}$ denote the normalized income distribution, with μ_{Z_m} indicating the average earnings in month m, and F_m the cumulative earnings distribution in m. Summing up the monthly distributions of a year gives the annual distribution, $Z_0 = \sum_{m=1}^{M} Z_m$. Dividing the annual distribution by average annual earnings gives the normalized distribution, $Y_0 = Z_0 / \mu_{Z_0}$. Finally, let F_0 denote the cumulative distribution function of annual earnings.

The Gini index can be computed from the covariance between the normalized income distribution and the cumulative distribution function. The Gini index for the earnings distribution in month m is $G_m = 2\operatorname{cov}\left(Y_m, F_m\left(Y_m\right)\right)$, and $G_0 = 2\operatorname{cov}\left(Y_0, F_0\left(Y_0\right)\right)$ is the Gini index for the annual distribution. In the sample, the empirical distribution substitutes for the theoretical one.[4]

The Gini coefficient has two asymmetric correlation coefficients associated with it. The Gini correlation coefficient, like the correlation coefficients of Pearson and Spearman, describes the dependence between two variables. The two Gini correlation coefficient from two distributions amongst themselves, say months m and n, are $\Gamma_{mn} = \frac{\operatorname{cov}\left(Y_m, F_n\left(Y_n\right)\right)}{\operatorname{cov}\left(Y_m, F_m\left(Y_m\right)\right)}$ and $\Gamma_{nm} = \frac{\operatorname{cov}\left(Y_n, F_m\left(Y_m\right)\right)}{\operatorname{cov}\left(Y_n, F_n\left(Y_n\right)\right)}$.

[4] In an equal probability sample, the rank of the variable is used to estimate the cumulative distribution function.

The range of the Gini correlation coefficient is $[-1,1]$. The Gini correlation coefficient has a number of important properties for our analysis.[5] If the distributions for two periods m and n are independent then $\Gamma_{mn} = 0$. If the rankings of observation units in the two distributions coincide (totally revert) then $\Gamma_{mn} = \Gamma_{nm} = 1$ ($\Gamma_{mn} = \Gamma_{nm} = -1$). Finally, the difference between the two Gini correlations is denoted by $D_{mn} = \Gamma_{mn} - \Gamma_{nm}$, $n,m = 0,...,M$. If the distributions are symmetric then the two correlation coefficients Γ_{mn} and Γ_{nm} are equal: $D_{mn} = \Gamma_{mn} - \Gamma_{nm} = 0$. For the two Gini correlations to be equal, a sufficient condition is that the variables Z_m and Z_n are exchangeable up to a linear transformation.

For the relationships between annual inequality and subannual inequality and mobility, the following identity holds (see Yitzhaki and Wodon, 2004, Proposition 1),

$$(1a) \quad G_0^2 = \sum_{m=1}^{M} \alpha_m^2 G_m^2 + \sum_{m=1} \sum_{m \neq n} \alpha_m \alpha_n G_m G_n \Gamma_{mn} + G_0 \sum_{m=1}^{M} \alpha_m D_{m0} G_m \,.$$

In Equation 1a, α_m denotes the sum of earnings in month m relative to the sum of annual earnings. Provided that all the pairwise Gini correlation coefficients are equal, $\Gamma_{mn} = \Gamma_{nm}$ (the Gini correlation coefficients for any two monthly distributions and for each monthly distribution with the annual distribution), then equation 1a simplifies to a quadratic two-component decomposition of the form,

$$(1b) \quad G_0^2 = \sum_{m=1}^{M} \alpha_m^2 G_m^2 + 2\sum_{m=1} \sum_{m<n} \alpha_m \alpha_n G_m G_n \Gamma_{mn} \,.$$

Equation 1b is identical in structure to the decomposition of the coefficient of variation. The first component, $\sum_{m=1}^{M} \alpha_m^2 G_m^2$, is the sum of the squared monthly Gini indices weighted by squared monthly earnings shares. The second component, $2\sum_{m=1} \sum_{m<n} \alpha_m \alpha_n G_m G_n \Gamma_{mn}$, is twice the sum of the

[5] Further information on the properties of the Gini correlation coefficient is provided in Schechtman and Yitzhaki (1987, 1999). In general, the properties are a mixture of Pearson's and Spearman's correlation coefficients.

product of the following factors: the two monthly earnings shares for m and n, the two Gini indices for m and n, and the Gini correlation coefficient for m and n.

From the Equations 1a and 1b it becomes evident that if the Gini correlation indices are not equal, two complications arise. While the first right-side component in Equations 1a and 1b is the same, $S_1 = \sum_{m=1}^{M} \alpha_m^2 G_m^2$, the second component in Equation 1a, $S_2 = \sum_{m=1} \sum_{m \neq n} \alpha_m \alpha_n G_m G_n \Gamma_{mn}$, embodies a larger number of summands compared with Equation 1b. Further, the additional component in Equation 1a, $S_3 = G_0 \sum_{m=1}^{M} \alpha_m D_{m0} G_m$, needs to be taken into account. This third component represents the differences between the two Gini correlations defined between each monthly distribution and the yearly distribution, $D_{m0}, m = 1,...,12$. It shows whether the monthly and the annual distribution are symmetric, which would mean that they belong to the same family of distributions.

2.2 Statistical inference

We apply the jackknife to provide confidence intervals for all the elements of the identity in Equation 1a. The basic idea behind the jackknife is that it systematically recomputes a statistic leaving one observation at a time out of the data. From the jackknife estimate of a statistic, an estimate of the variance of the statistic can be derived (see Wolter, 1985).

Let M_{-k} denote the jackknife statistic when the k^{th} observation is taken out of the sample. Let M denote the average of M_{-k}. Further, let w_k denote the frequency weight of k. Then the jackknife variance estimator is,

$$(2) \quad v = \frac{N^w - 1}{N^w} \sum_k w_k \left(M_{-k} - M \right)^2 ,$$

where $N^w = \sum_k w_k$ denotes the weighted number of observations.

To reduce the computational burden related to the computation of the Gini indices and Gini correlation coefficients, we implemented procedures that calculate, regardless of sample size, jackknife estimates with only a few passes through the data (for detail see Yitzhaki, 1991; Karagiannis and Kovacevic, 2000). An alternative to the jackknife is the bootstrap. Leaving theoretical differences aside, the main practical difference between the bootstrap and the jackknife is that, when repeated on the same data, only the jackknife yields exactly the same result each time. For this reason we use the jackknife.

3 Data base and working sample

Our empirical analysis builds on the Insurance Account Sample for the year 2006 (IAS 2006, "Versicherungskontenstichprobe"), i.e., administrative data on employment records provided by the Data Center of the German Pension Insurance (Forschungsdatenzentrum der Rentenversicherung). The scientific use file IAS 2006 details information on the employment records of 60,304 individuals.

IAS is a survivor sample: the population is pension insurants alive in 2006. For this population, IAS provides biographical information on various characteristics relevant for the pension insurance. Sample weights are provided to generate results that are representative for the pension insurants ages 15 to 67 and still alive in 2006. These sampling weights are always used in the subsequent inequality analysis. Most importantly, for every month of the employment biography of an insurant, the individual earnings subject to social security can be derived from the documented monthly pension contributions.[6] Hence, information on earnings is available for an accounting period of one month. By adding up the twelve monthly earnings over a year, we obtain the annual earnings. The distributions of monthly and annual earnings allow for a systematic investigation of the IAP-distribution nexus under ceteris paribus conditions.

Five issues should be noticed. First, IAS does not provide information on other income sources, on family background, or on income tax burdens.[7] For this reason, our empirical analysis will rely on nominal gross individual earnings. Note, however, that the Gini coefficient is a relative

[6] In Germany, individual pension contributions are directly related to individual earnings up to an assessment ceiling.
[7] Marital status is unknown, and children are recorded only if they determine the individual pension entitlement. As child-care periods are usually credited only for female insurants, children of male insurants typically cannot be observed.

measure, so that inflation and/or economic growth do not affect it. Second, not all employees are mandatorily insured in the German Statutory Pension Insurance. The two most important groups that are not insured and, therefore, not included in the database, are civil servants and the self-employed. Third, social security assessment ceilings[8] exist, so that earnings information in IAS is top coded. Over the observation period, top coding affects slightly more than six percent of the observations in our working sample. Over our observation period, the fraction of top-coded observations does not follow a systematic inter-temporal pattern period. It ranges between 4.6 percent in 2003 and 7.6 percent in 2001.[9] Top-coding implies that our inequality estimates tend to underestimate the actual level of earnings inequality. Fourth, assessment ceilings have risen over time, i.e., from EUR 3,323 per month in 1991 to EUR 5,250 per month in 2006. Because this affects the threshold for the top coding, inter-temporal changes of distributional indices should be interpreted with adequate care. The fifth issue to be taken into account is that the sample is a sample of survivors. Hence, the sample is not useful for describing overall earnings inequality in Germany. However, it is useful for our purposes because we limited the sample to a certain age group.

Our analysis starts right after German reunification in 1991. To immunize the statistics from other blurring factors, we focus on a rather homogeneous sample, i.e., males and females age 30 to 50 in the western German states whose social status in all twelve months of a particular year is: employed, marginally employed, or unemployed. Accordingly, those not participating in the labor market or who are employed but not subject to compulsory insurance are discarded from the database. To secure a consistent implementation of the Gini decomposition, the working sample is restricted to insurants that can be tracked over all twelve months of the year under investigation.

Descriptive (non-weighted) statistics on the working sample's size and composition are summarized in Table 1. Altogether, the working sample consists of about 13,000 to 16,000 insurants per year. By construction, the age composition of the sample is quite stable over time, with the average age always being slightly below 40 years. Nominal gross individual earnings

[8] For further details on levels of assessment ceilings see Appendix B in Schröder (2012).

[9] Fractions relate to the non-weighted working sample. Weighted fractions hardly differ. For a database constructed from the IAS waves from 2005 to 2008, Bönke et al. (2011) report a higher number of top-coded units of about ten percent. However, both their database and their working sample are different from ours.

increased substantially over the observation period, from EUR 22,851 in 1991 to EUR 31,053 in 2006. The last three columns in the table provide information on the employment status of the insurants. Specifically, the columns provide the average number of months in a particular year spent as employed, unemployed, and marginally employed.[10] As an example, in year 2006, the average insurant in our working sample was fully employed for 11.562 months, unemployed for 0.351 months, and marginally employed for 0.087 months.

Table 1 about here

4 Results from the Gini decomposition

All the results from the Gini decomposition consider IAS sample weights. Because of the construction of the working sample and because IAS is a survivor sample, our results are representative for employed, marginally employed, and unemployed, prime-age, West German pension insurants alive in 2006. Insurants who died before 2006 are not included in any of the 1991-2006 earnings distributions.

4.1 The central elements of the decomposition

Equation 1a decomposes annual inequality, captured by the squared annual Gini, into three components: $S_1 = \sum_{m=1}^{M} \alpha_m^2 G_m^2$, $S_2 = \sum_{m=1} \sum_{m \neq n} \alpha_m \alpha_n G_m G_n \Gamma_{mn}$, and $S_3 = G_0 \sum_{m=1}^{M} \alpha_m D_{m0} G_m$. Remember that the first component is the sum of the squared monthly Gini indices weighted by squared monthly earnings shares. The second component is the sum of the products of two monthly earnings shares, two monthly Gini indices, and the Gini correlation coefficients of the two months. The third component represents the differences between two Gini correlations defined between each monthly and the yearly distribution.

Figure 1 depicts the trends of these three components together with the squared Gini index derived from the annual earnings distribution. Each measure is presented in a separate graph.

[10] Due to changes in the German social security code, marginal employees are part of the database only as of 1999.

Vertical lines represent the 95 percent jackknife confidence intervals, and crosses indicate the point estimates from the full working sample.

Figure 1 about here

The upper left graph provides the squared Gini indices from the annual earnings distribution. The graph shows a prominent rise in inequality in the annual earnings distribution over the observation period. The squared Gini index from the annual distribution (point estimate) increased by almost 61 percent, from 0.056 in 1991 to 0.090 in 2006. As the upper bound of the jackknife confidence intervals in the early 1990s is substantially lower than the lower bound in the 2000s, the rise is statistically significant. Most of the rise occurred in the most recent years. In our database, inter-temporal changes in inequality should be interpreted with care, because assessment ceilings create downward bias in the inequality indices. However, assessment ceilings, due to stagnating earnings, hardly differ in the later years. As a result, for relative measures like the Gini index, the bias should not change systematically. As the bias should not change systematically, the recent rise in measured inequality should mirror a real-world fact.[11]

The upper right graph and the lower left graph show the components $S_1 = \sum_{m=1}^{M} \alpha_m^2 G_m^2$ and $S_2 = \sum_{m=1} \sum_{m \neq n} \alpha_m \alpha_n G_m G_n \Gamma_{mn}$. Both sums closely follow the pattern of the squared annual Gini. However, the first component, S_1, is only about a tenth of the second component, S_2. Indeed, as we will show, S_2 plays the crucial role in the recent rise in annual earnings inequality in Germany. Finally, the lower right graph provides the third component, $S_3 = G_0 \sum_{m=1}^{M} \alpha_m D_{m0} G_m$. The component S_3 turns out to be quantitatively small and exhibits no systematic variation over the observation period, with point estimates fluctuating around -0.0002. As they are close to zero, the

[11] One way to evaluate the effect of ceilings change on inequality is to identify the percentage of employees above the ceiling. This percentage did not change much since reunification. Another way is to impute earnings, assuming that the upper part of the earnings distribution follows a particular distribution, e.g., the Pareto-distribution.

distributions are exchangeable up to a linear transformation, and are about symmetric with each other. This is an interesting preliminary finding, implying that as an approximation the component S_3 can be ignored, thus suggesting that the quadratic two-component decomposition (Equation 1b) can serve as a good approximation. We provide formal tests of the symmetry of Gini correlation coefficients in Section 4.3.

The stability of S_3 is a bit surprising, given the pronounced rise in the annual Gini index, G_0. As we shall show, its stability results from the inter-temporal decline of the absolute differences of the Gini correlations between subannual and annual earnings distribution, $|D_{m0}|$. Their decline "balances" the rise in the monthly Gini indices, G_m. To better understand the mechanics underlying these patterns, the next subsection provides the disaggregated components of the Gini decomposition.

4.2 The subannual indices

The rise of inequality in the annual earnings distribution in general and of the component $S_2 = \sum_{m=1}\sum_{m\neq n}\alpha_m\alpha_n G_m G_n \Gamma_{mn}$ in particular may be due to changes in levels and interactions of three distributional statistics: the monthly shares in total annual earnings, α_m, the monthly Gini indices, G_m, and the Gini correlation coefficients from monthly distributions amongst themselves, Γ_{mn}. For example, it could be that earnings shares related to months with relatively low inequality have decreased, and/or that the Gini correlation coefficients, Γ_{mn}, have increased, and/or that the monthly distributions have become distributed more unequally.

We begin the analysis by asking whether the increase in inequality in the annual earnings distribution and in S_2 resulted from changes in the monthly shares of total annual earnings. As seen in Figure 2, the answer is no. The figure provides point estimates and confidence intervals of the monthly earnings shares. Horizontal lines provide a fictitious scenario where monthly shares of annual earnings are identical in all months of a year. Not surprisingly, the figure tells the following consistent story over the entire observation period. Monthly earnings shares are slightly lower in the winter than in the summer months (due to seasonal effects in the labor market) and are directly related to the number of (working) days, as reflected by the exceptionally low earnings share in February. As both patterns are documented consistently over the entire observation period, the rise in annual inequality and in S_2 must be due to rising inequality in the monthly distributions, and/or higher Gini correlation coefficients from monthly distributions amongst themselves.

Figure 2 about here

For this reason, we turn next to the Gini indices derived from the monthly earnings distributions. Results (95 percent confidence intervals and point estimates) are assembled in Figure 3 in twelve graphs from January ("Month 1") to December ("Month 12"). Further, to allow for a quantitative assessment of the difference between monthly and annual inequality, Gini indices from the annual distribution are shown in gray.[12] The first finding is that the monthly indices closely track the annual indices. A period of stability until 2000 is followed by a period with rising earnings inequality. The second finding is that the monthly Gini indices are typically higher than their annual counterpart in the same year. This is because an extension of the accounting period from a month to a year smoothes subannual earnings fluctuations (Shorrocks, 1978a). The difference in monthly and annual inequality is most pronounced in the winter months December to March. One-time payments in December (Christmas bonuses) and seasonal changes in the labor market are the most plausible explanation for this. For example, unemployment rates in the winter months January to March, in 2006 fluctuated around twelve percent, and thus were about two percentage points higher than in the other months. Moreover, the number of short-term contracts was also significantly higher in the winter than in the other seasons. In January, February, and March 2006, about 100,000 people had a short-term contract, in July 2006 the number was 50,000 people, and in November it was 36,000.[13]

[12] Results from the annual distribution are shifted slightly to the left to be visually distinguishable from the monthly estimates.
[13] See the Federal Statistical Office at https://www.destatis.de/EN/Homepage.html for details.

Figure 3 about here

Finally, we turn to the Gini correlation coefficients from monthly distributions amongst themselves, Γ_{mn}, summarized in Figure 4. Altogether, eleven correlations can be computed for every month, resulting in a total of $11 \times 12 = 132$ correlations per year. However, we will confine the presentation to the correlations between the Gini index for January ($m = 1$) and the Gini indices for February to December, i.e., Γ_{1n}, as the same patterns hold for the remaining months (see Figures A1-A11 in the Appendix). Most noticeably, the monthly Gini correlation coefficients, Γ_{mn}, are all rather high, being above 0.88. As explained in Section 2, a correlation of 1.0 indicates the absence of rank exchange mobility, and thus the Gini correlation coefficients suggest that the subannual rank exchange mobility within the time span of a year is rather low.

Looking at the Gini correlation coefficients from the monthly distributions amongst themselves in more detail, it is evident that they become smaller as the time span between the two months under investigation is extended. In 2006, for example, the Gini correlation coefficients (point estimate) for January and February is $\Gamma_{12} = 0.999$, for January and July it is $\Gamma_{17} = 0.957$, and for January and December it is $\Gamma_{112} = 0.942$. An interesting finding is that the Gini correlation coefficients from the monthly distributions are of similar size until 2004 but are markedly higher for 2005 and 2006. Accordingly, rank exchange mobility has recently decreased.[14] This finding is interesting in its own right, but its implications are more general. As pointed out in Wodon and Yitzhaki (2003, p. 4), if Gini correlation coefficients "are relatively stable over time […], we may be able to predict the impact of the accounting period on inequality in quite general settings." Our finding, however, indicates that Gini correlation coefficients are not stable, shedding doubts on the predictability of the quantitative impact of the IAP on inequality measures.

However, if the Gini correlation coefficients increase over time, we should expect a lower bias due to differences in accounting periods. To convey the basic argument behind this conjecture,

[14] Assessment ceilings hardly changed in the period 2004 to 2006. These changes cannot explain the prominent rise in rank exchange mobility.

think about a limited case. Assume that the distributions are symmetric, so that the third term from Equation 1a can be ignored, and also $G_m = C$ for $m = 1,...,M$, and $\Gamma_{mn} = B \; \forall m,n$. Then from Equation 1a it follows that $G_0^2 = C^2 \left\{ \sum_{m=1}^{M} \alpha_m^2 + B \sum_{m=1} \sum_{m \neq n} \alpha_m \alpha_n \right\}$. Further, assume that $B = 1$. It follows that $G_0^2 = C^2 \left\{ \sum_{m=1}^{M} \alpha_m \sum_{n=1}^{M} \alpha_n \right\} = C^2$: the maximum value of the Gini index from the yearly distribution is equal to the Gini index from the monthly distributions. Otherwise, because the earnings shares, α, are positive, it is easy to see that an increase in B increases the yearly Gini index so that the yearly Gini index approaches the monthly Gini index.

For an empirical evaluation of the size of the bias, it is interesting to see how the average of the Gini correlation coefficients has evolved over the observation period. For our working sample, the average of the Gini correlation coefficients has increased over the observation period: from 0.943 in 1991 to 0.956 in 1995, 0.960 in 2000, and 0.971 in 2006. This increase means that the bias due to differences in accounting periods has become lower. A plausible explanation for the increase in monthly correlation coefficients (and Gini indices) over time is that employees tend to want to stay at their jobs when labor markets become unstable.

Figure 4 about here

4.3 The symmetry tests for Gini correlation coefficients

Equation 1a decomposes the square of the annual Gini index into three components. Whether the simpler quadratic two-component decomposition, Equation 1b, can serve as an adequate first-order approximation for applied inequality researchers depends on whether the Gini correlation coefficients are symmetric or asymmetric. This section tests the symmetry of these correlations.

Following Finkel et al. (2006), our test statistic is the difference between two Gini correlation coefficients divided by its jackknife standard error. For a large number of observations, the statistic's critical value at the 95 percent confidence level is 1.96. The tests can be performed for Gini correlation coefficients for the annual distribution and the monthly distributions,

$t_{m0} = |D_{m0}| / SE(D_{m0})$, and also for the Gini correlations for any two monthly distributions amongst themselves, $t_{mn} = |D_{mn}| / SE(D_{mn})$. Since $|D_{mn}| = |D_{nm}|$, the test statistics for two monthly distributions amongst themselves, t_{mn} and t_{nm}, coincide. Altogether, 78 hypotheses need to be tested for every year: $M \cdot (M-1)/2 = 66$ tests for the Gini correlation coefficients for the monthly distributions amongst themselves, plus twelve test statistics for the correlations for the monthly distributions and the annual distribution.

Due to the multiplicity of comparisons, there is also the issue of "family-wise error rates." Given a set probability $\alpha = 0.05$ of a Type-1 error (i.e., incorrect rejection of a true null hypothesis, a false positive), about one out of every twenty such tests will show a false positive. Per year, we perform twelve tests for the Gini correlation coefficients for the monthly distributions and the annual distribution, and 66 tests for the Gini correlation coefficients for the monthly distributions amongst themselves. Hence, we expect 0.6 of the tests or 3.3 tests to be declared as significant if we use $\alpha = 0.05$ for each test. This is the problem of family-wise error rates. To counteract this problem, several methods have been suggested in the literature (for a review and recent advancements see Romano and Wolf, 2005).

We apply the Holm-Bonferroni correction (Holm, 1979), a sequential correction procedure, which does not require tests to be independent. The correction relies on the twelve (or 66) p-values for the monthly-annual (or monthly-monthly) tests. Let X denote the number of tests being performed, and let the p-values for the $x = 1, ..., X$ hypotheses being tested, ordered from smallest to largest, be $p(1) \leq p(2) \leq ... \leq p(X)$. Further, let $H(x)$ be the hypothesis associated with the p-value for x. To control for the family-wide false positive value (FWER), π, proceed as follows:

1. If $p(1) > \pi / X$, accept all the x hypotheses (none are significant).
2. If $p(1) \leq \pi / X$, reject $H(1)$ [$H(1)$ is declared significant], and consider $H(2)$.
3. If $p(2) > \pi / (X - 2 + 1)$, accept $H(x)$ (for $x \geq 2$).
4. If $p(2) \leq \pi / (X - 2 + 1)$, reject $H(2)$ and consider $H(3)$.

5. Proceed with the hypotheses until the first x such that $p(x) > \pi/(X - x + 1)$.

Table 2 provides the results in two portions. The first portion (Columns 1 and 2) shows the results from the sequential hypothesis testing using the Holm-Bonferroni correction. The first column relates to the testing of the differences of the Gini correlations for the annual distribution and the monthly distributions, $D_{m0} = \Gamma_{m0} - \Gamma_{0m}$. The second column relates to the testing of differences of the Gini correlation coefficients for the monthly distributions amongst themselves, $D_{mn} = \Gamma_{mn} - \Gamma_{nm}$ ($m, n = 1, \dots, 12$ and $m \neq n$). In each of the two columns, we report the fraction of significant tests under the Holm-Bonferroni correction. The second portion (Columns 3 to 8) is descriptive information (min, mean, max) on the two differences D_{m0} and D_{mn}.

Table 2 about here

Regarding the differences of the Gini correlations for the annual distribution and the monthly distributions, $D_{m0} = \Gamma_{m0} - \Gamma_{0m}$, all the tests are significant. Regarding the Gini correlation coefficients for the monthly distributions amongst themselves, the fraction of significant tests ranges between 60.61 percent in 1993 and 83.33 percent in 1997. However, it turns out that all the differences are always small in quantitative terms. The differences D_{m0} are small, have a negative sign, and range between -0.0034 (1991) and -0.0016 (2006). The differences D_{mn} are small, have an ambiguous sign, and range between -0.0018 (1998) and 0.0029 (1991).

Small differences between Gini correlations are also reported in Finkel et al. (2006), who therefore argued that the simpler two-component decomposition may be viewed as a helpful first-order approximation "for any practical purpose" (p. 158). However, the consistency of estimates does not ensure that any conclusions are generalizable to other income concepts, time periods, or countries, given the lack of a theoretical foundation. Indeed, using Mexican data on individual earnings, Wodon and Yitzhaki (2003) found larger differences between Gini correlation coefficients, albeit without providing formal tests of significance of the differences.

5 Conclusions

This paper investigates the relationships between inequality in the annual earnings distribution, inequality in the underlying monthly distributions, and rank exchange mobility during the course of the year using the Gini decomposition approach outlined in Wodon and Yitzhaki (2003) and in Yitzhaki and Wodon (2004). Employment histories of German pension insurants for the period 1991 to 2006 were used as the database.

Three findings are worth mentioning from a methodological perspective. First, the IAP does not impact the general inter-temporal pattern of inequality. Gini indices from both the annual and monthly earnings distributions indicate that inequality has increased over time.[15] The explanation is that the increase in monthly inequality was accompanied by an increase in correlations, so that both factors were in the same direction. Second, Gini indices from the monthly distributions are significantly higher than Gini indices from the annual distributions. This finding questions the validity of comparative distributional studies that build on data that do not use a uniform accounting period. Third, Gini correlation coefficients of the monthly distributions amongst themselves change over time. This makes it difficult to make a general prediction of the quantitative impact of the accounting period on inequality.

The good news for practitioners is that the difference in Gini correlations between monthly and yearly distributions is small for all practical purposes, implying that the additional parameters that the Gini decomposition includes may be ignored, and that the simpler formula that is used to decompose the coefficient of variation provides a reasonable approximation of the Gini decomposition.

Acknowledgements

We would like to thank the Editor of the Journal and two anonymous Reviewers for helpful comments. We would like to thank Ralf Himmelreicher of the Research Data Centre of the German Pension Insurance for guidance through the process of data preparation. Carsten Schröder would like to thank the Kiel Institute for the World Economy for its kind hospitality during a research visit from April to September 2012. The usual disclaimer applies.

[15] Because earnings in IAS are top coded, our inequality estimates are probably biased downwards when applied to the German labor market.

References

Atkinson, A.B.: The Measurement of Economic Mobility, Social Justice and Public Policy, Wheatsheaf Books Ltd., Sussex (1983)

Atkinson, A.B., and Bourguignon, F.: Empirical Studies of Earnings Mobility, Harwood: Chur (1992)

Böheim, R., and Jenkins, S.P.: A Comparison of Current and Annual Measures of Income in the British Household Panel Survey, Journal of Official Statistics, **22** 733–758 (2006)

Böheim, R., and Jenkins, S.P.: Do Current Income and Annual Income Measures Provide Different Pictures of Britain's Income Distribution, DIW Discussion Papers, **214** (2000)

Bönke, T., Corneo, G., and Lüthen, H.: Lifetime Earnings Inequality in Germany, Institute for the Study of Labor, **6020** (2011)

Canberra Group: Expert Group on Household Income Statistics, Final Report and Recommendations, Ottawa (2001)

Cantó, O., Gradin, C., and Del Río, C.: Poverty Statics and Dynamics: Does the Accounting Period Matter?, International Journal of Social Welfare **15**, 209-218 (2006)

Chesher, A., and Schluter, C.: Welfare Measurement and Measurement Error, Review of Economic Studies **69**, 357-378 (2002)

Dardanoni, V.: Measurement of Social Mobility, Journal of Economic Theory **61**, 372-394 (1993)

Detlefsen, L.: Earnings Inequality - Does the Accounting Period Matter?, Schmollers Jahrbuch **132**, 297-321 (2012)

Finkel, Y., Artsev, Y., and Yitzhaki, S.: Inequality Measurement and the Time Structure of Household Income in Israel, Journal of Economic Inequality **4**, 153-179 (2006)

Gibson, J., Huang, J., and Rozelle, S.: Why is Income Inequality so Low in China compared to Other Countries? The Effect of Household Survey Methods, Economic Letters **71**, 329-333 (2001)

Holm, S.: A Simple Sequentially Rejective Multiple Test Procedure, Scandinavian Journal of Statistics **6**, 65–70 (1979)

Karagiannis, E., and Kovacevic, M.: A Method to Calculate the Jackknife Variance Estimator for the Gini Coefficient, Oxford Bulletin of Economics and Statistics **62**, 119-122 (2000)

King, M.A.: An Index of Inequality: With Application to Horizontal Equity and Social mobility, Econometrica **51**, 99–115 (1983)

Kühnen, C.: Das Stichprobenverfahren der Einkommens- und Verbrauchsstichprobe 1998, Methodenberichte, 1/2001, Wiesbaden: Statistisches Bundesamt (2001)

Lerman, R.I., and Yitzhaki, S.: A Note on the Calculation and Interpretation of the Gini index, Economics Letters **15**, 363-368 (1984)

Morris, N., and Preston, I.: Inequality, Poverty and the Redistribution of Income, Bulletin of Economic Research **38**, 277-344 (1986)

Nolan, B.: Income Distribution and the Macroeconomy, Cambridge University Press: Cambridge (1987).

Pyatt, G., Chen, C.N., and Fei, J.: The Distribution of Income by Factor Components, Quarterly Journal of Economics **95**, 451-473 (1980)

Ravallion, M. (1988): Expected Poverty under Risk-induced Welfare Variability, Economic Journal **98**, 1171–1182 (1987)

Romano, J.P., and Wolf, M.: Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing, Journal of the American Statistical Association **100**, 94-108 (2005)

Ruggles, P.: Drawing the Line: Alternative Poverty Measures and Their Implications for Public Policy, The Urban Institute Press: Washington DC (1990)

Schechtman, E., and Yitzhaki, S.: A Measure of Association Based on Gini's Mean Difference, Communications in Statistics Theory and Methods **A16,1**, 207-231 (1987)

Schechtman, E., and Yitzhaki, S.: On the Proper Bounds of the Gini Correlation, Economics Letters **63**, 133-138 (1999)

Schröder, C.: The Sensitivity of Distributional Measures to the Reference Period of Income, Kiel Working Papers, 1777, Kiel Institute for the World Economy (2012)

Shorrocks, A.J.: Income Inequality and Income Mobility, Journal of Economic Theory **19**, 376-393 (1978a)

Shorrocks, A.J.: Measurement of Mobility, Econometrica **46**, 1013–1024 (1978b)

Shorrocks, A.J.: Income Inequality and Income Mobility, Journal of Economic Theory **19**, 376–393 (1978c)

Shorrocks, A.J.: Income Stability in the United States, in N.A. Klevmarken and J.A. Lybeck (eds.), The Statics and Dynamics of Income, Tieto, Clevedon, Avon (1981)

Wodon, Q., and Yitzhaki, S.: Inequality and the Accounting Period, Economics Bulletin **4**, 1-8 (2003)

Wolter, K.: Introduction to Variance Estimation, Springer, New York (1985)

Yitzhaki, S., and Wodon, Q.: Mobility, Inequality and Horizontal Equity, Research on Economic Inequality **12**, 179-199 (2004)

Yitzhaki, S.: Calculating Jackknife Variance Estimators for Parameters of the Gini Method, Journal of Business and Economic Statistics **9**, 235-239 (1991)

Table 1. Description of the sample

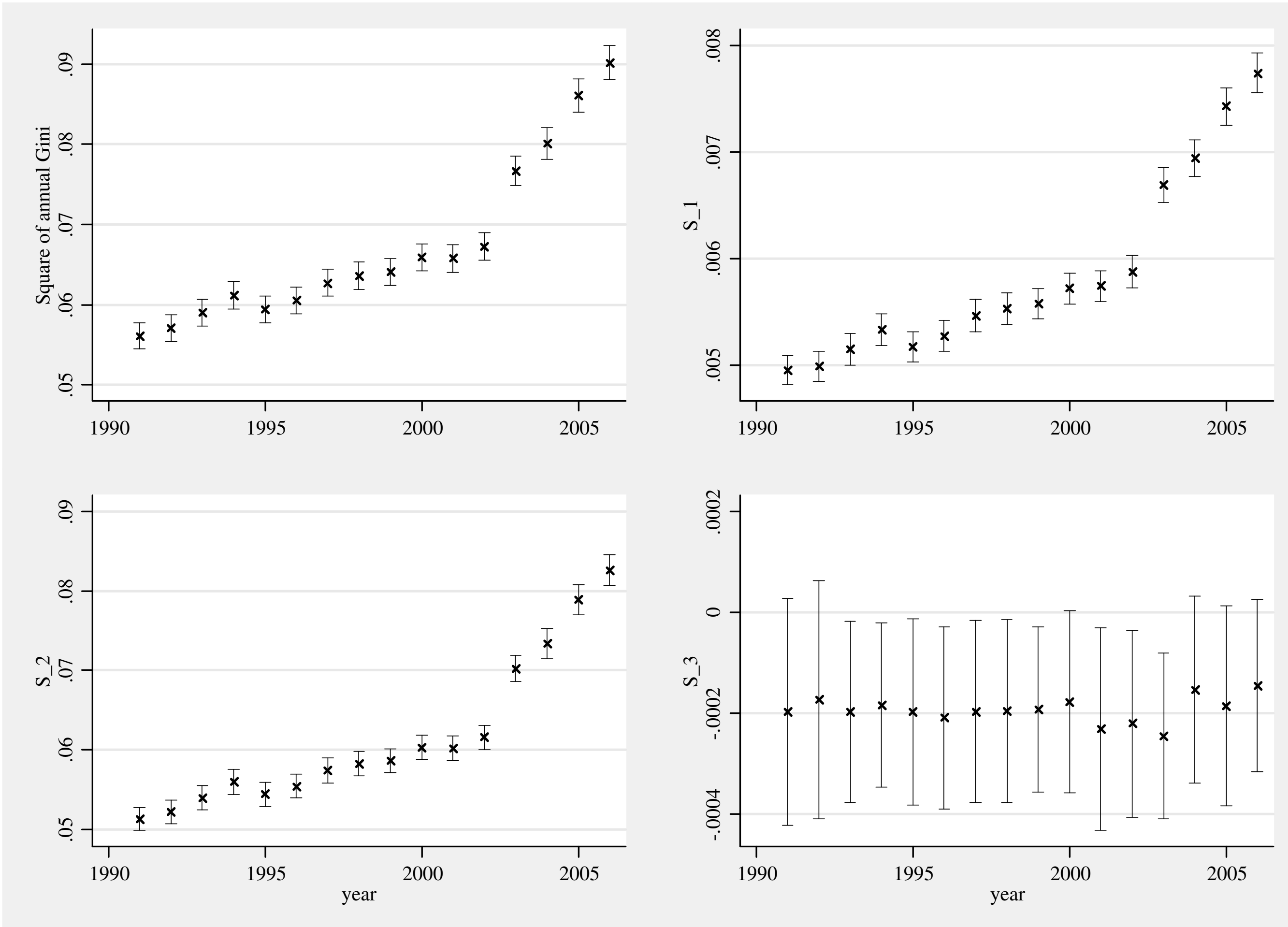
Year	Number of observations	Age	Earnings	Months spent in status of: unemployment	marginal employment	full employment
1991	12930	39.907	22850.502	0.249	0.000	11.751
1992	13050	39.798	24352.371	0.334	0.000	11.666
1993	13168	39.658	25446.639	0.387	0.000	11.613
1994	13132	39.582	26075.725	0.397	0.000	11.603
1995	13109	39.490	27065.273	0.364	0.000	11.636
1996	13238	39.515	27648.961	0.404	0.000	11.596
1997	13328	39.494	27855.381	0.447	0.000	11.553
1998	13735	39.498	28247.668	0.404	0.000	11.596
1999	14135	39.529	28755.549	0.419	0.014	11.567
2000	14717	39.514	28895.100	0.381	0.077	11.543
2001	15080	39.506	29696.709	0.373	0.070	11.557
2002	15316	39.500	30096.580	0.428	0.054	11.518
2003	15486	39.538	30900.396	0.453	0.063	11.484
2004	15449	39.583	31197.033	0.449	0.066	11.485
2005	15931	39.583	30795.566	0.411	0.074	11.515
2006	16057	39.566	31053.182	0.351	0.087	11.562

Note. Database is IAS 2006. Own calculations.

Table 2. Test statistics

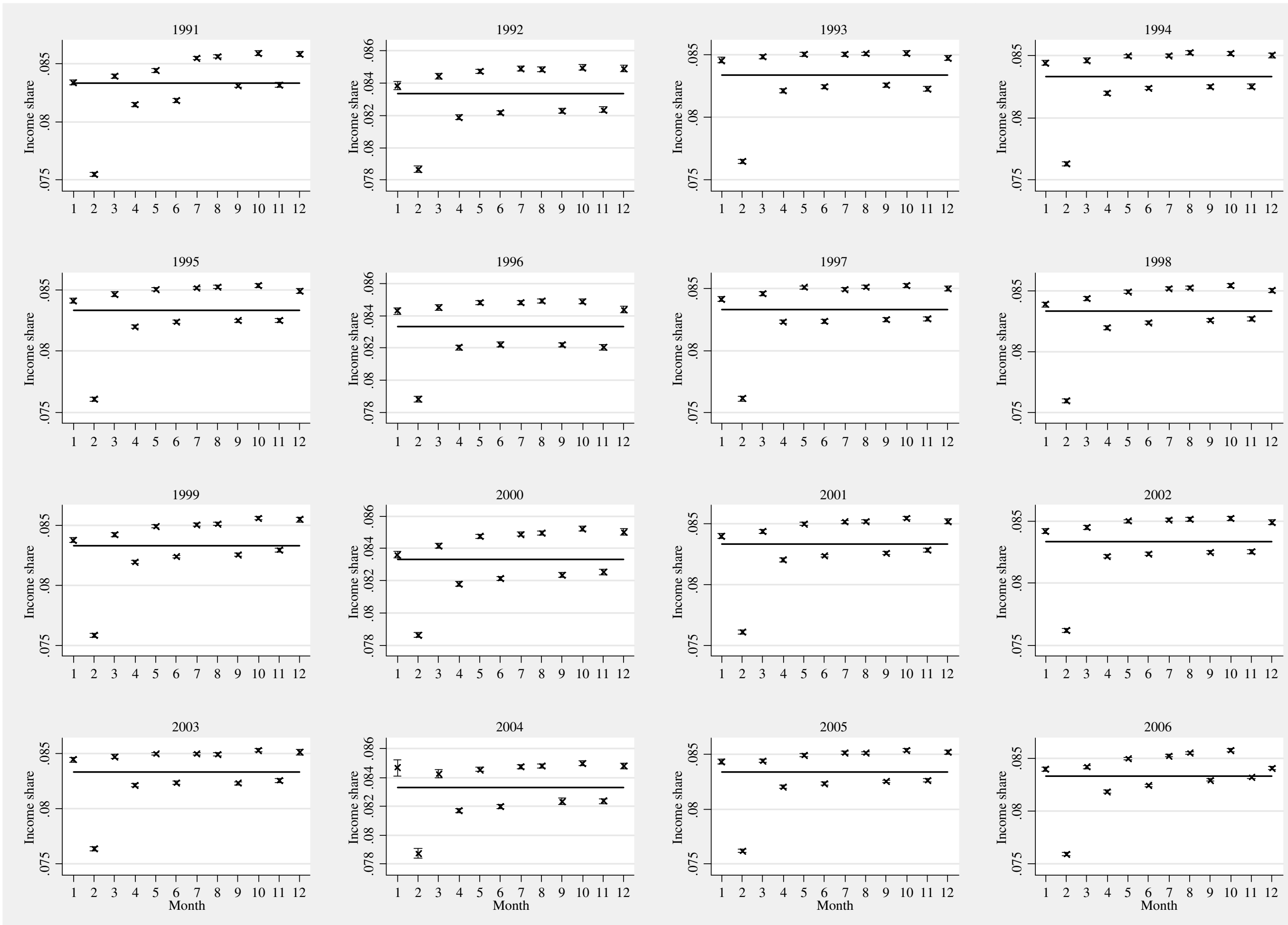
Year	Share of significant tests under Holm-Bonferroni correction (in %)		D_{m0}			D_{mn}		
	$\lvert D_{m0}\rvert \neq 0$	$\lvert D_{mn}\rvert \neq 0$	Min	mean	max	min	mean	max
1991	100	75.76	-0.006	-0.003	-0.002	-0.006	0.003	0.010
1992	100	63.64	-0.007	-0.003	-0.001	-0.006	0.001	0.005
1993	100	60.61	-0.006	-0.003	-0.002	-0.002	0.002	0.006
1994	100	65.15	-0.005	-0.003	-0.001	-0.003	0.000	0.005
1995	100	68.18	-0.006	-0.003	-0.001	-0.007	-0.001	0.004
1996	100	68.18	-0.006	-0.003	-0.002	-0.006	0.000	0.005
1997	100	83.33	-0.006	-0.003	-0.001	-0.007	0.000	0.005
1998	100	78.79	-0.006	-0.003	-0.001	-0.008	-0.002	0.005
1999	100	68.18	-0.006	-0.003	-0.001	-0.007	-0.001	0.005
2000	100	72.73	-0.006	-0.003	-0.001	-0.006	-0.001	0.005
2001	100	78.79	-0.008	-0.003	-0.001	-0.006	0.001	0.010
2002	100	63.64	-0.007	-0.003	-0.002	-0.001	0.002	0.008
2003	100	81.82	-0.006	-0.003	-0.001	-0.003	0.002	0.008
2004	100	66.67	-0.005	-0.002	-0.001	-0.009	-0.001	0.005
2005	100	74.24	-0.004	-0.002	-0.001	-0.003	0.000	0.004
2006	100	65.15	-0.003	-0.002	-0.001	-0.004	-0.001	0.003

Note. Database is IAS 2006. Own calculations.



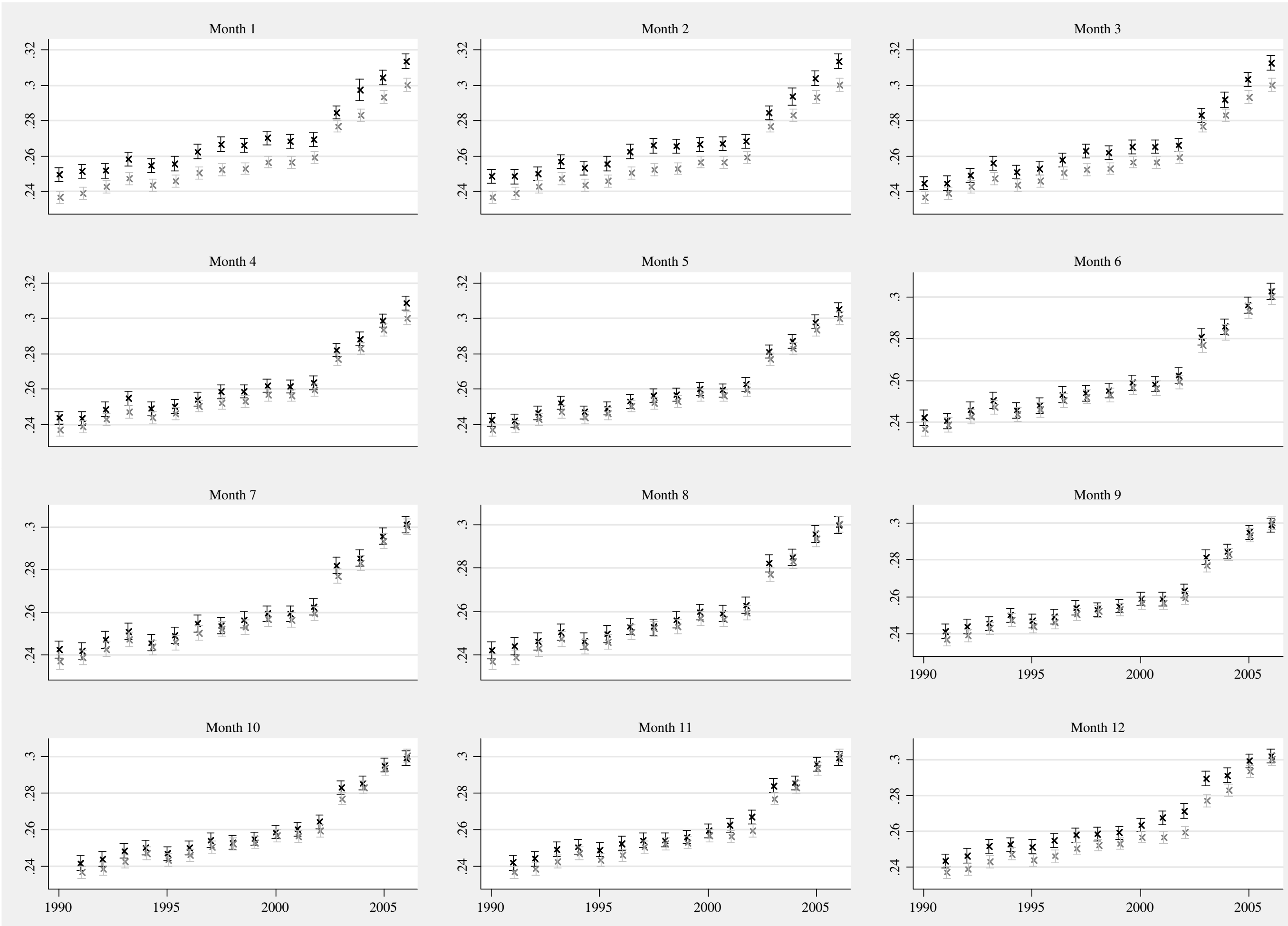
Note. Database is IAS 2006. Own calculations.

Figure 1. Estimates from the decomposition



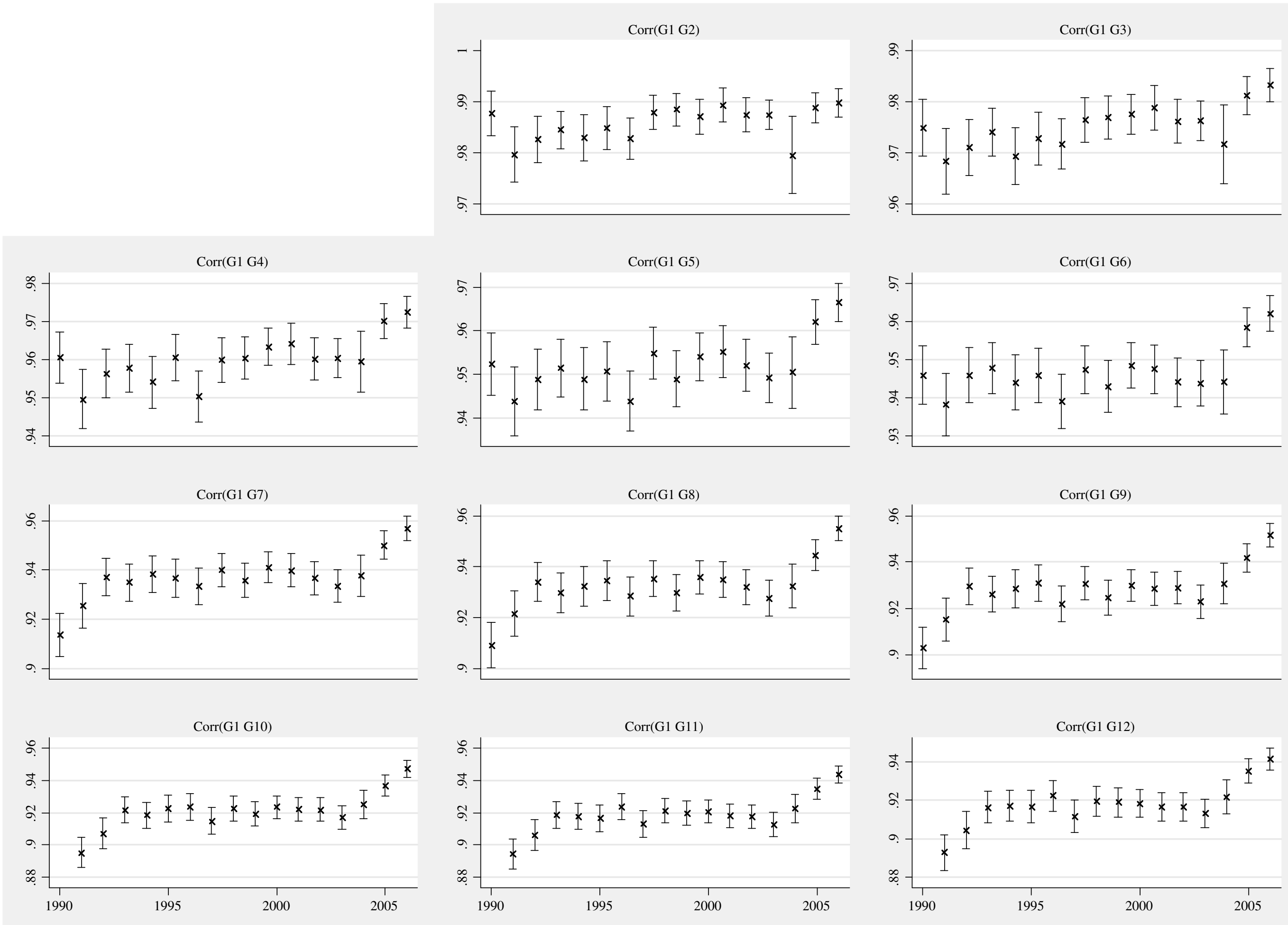
Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; "x": point estimate.

Figure 2. Periodic earnings shares



Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; “x”: point estimate. Monthly estimates appear in black; annual estimates in grey.

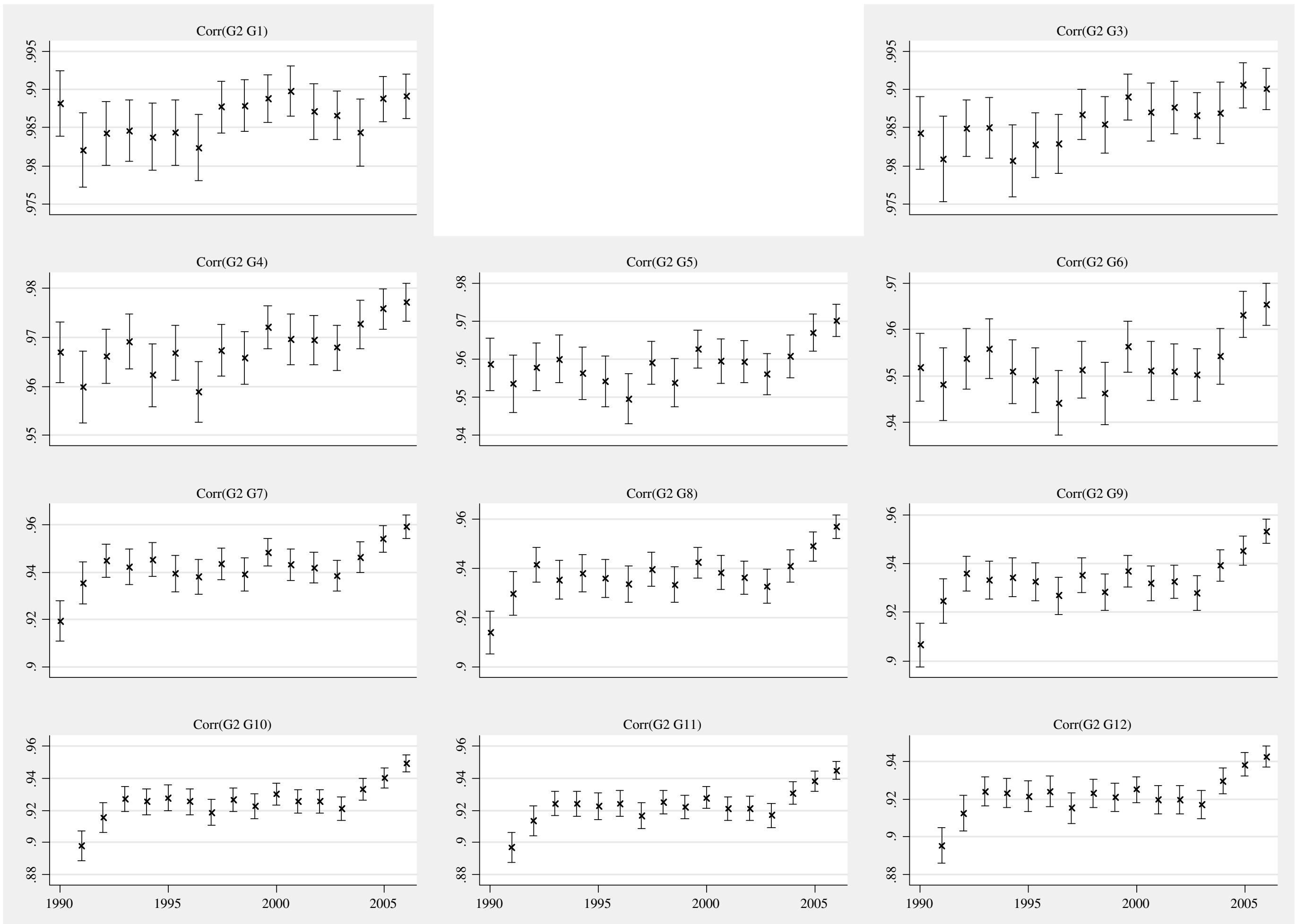
Figure 3. Periodic and annual Gini coefficients



Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; "x": point estimate.

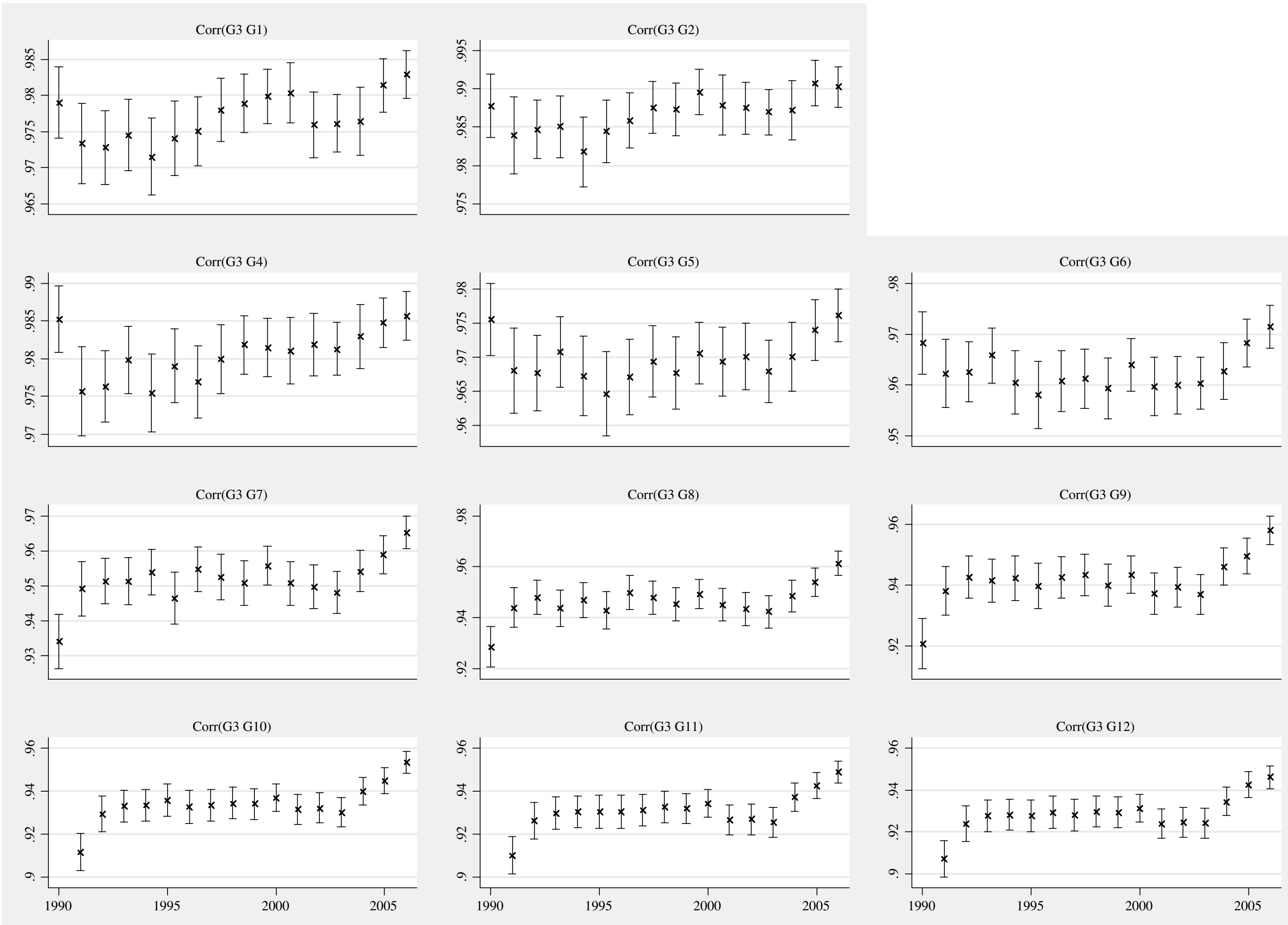
Figure 4. Gini correlations

Appendix: intended for online use only



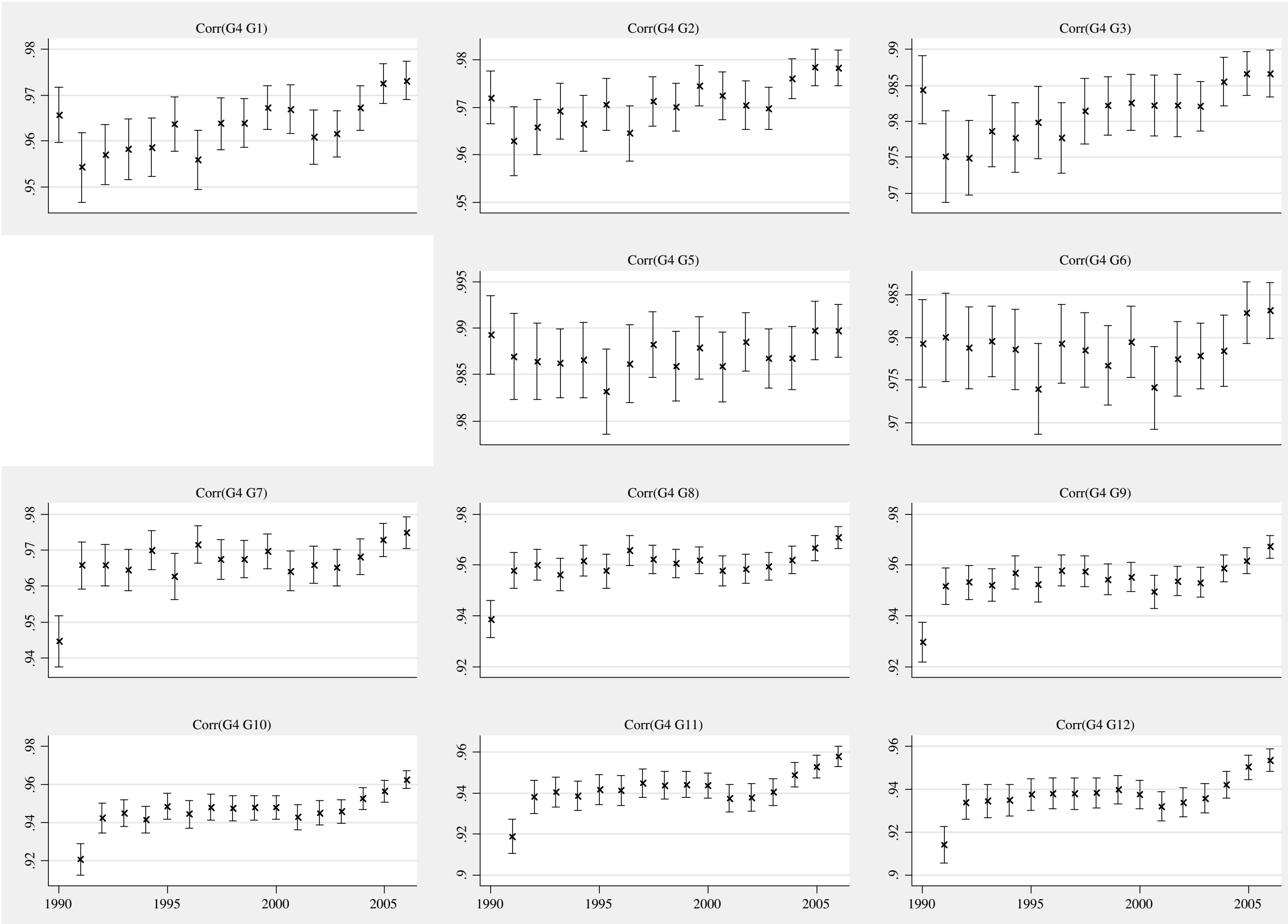
Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; “x”: point estimate.

Figure A1. Gini correlations for February with other months of the year



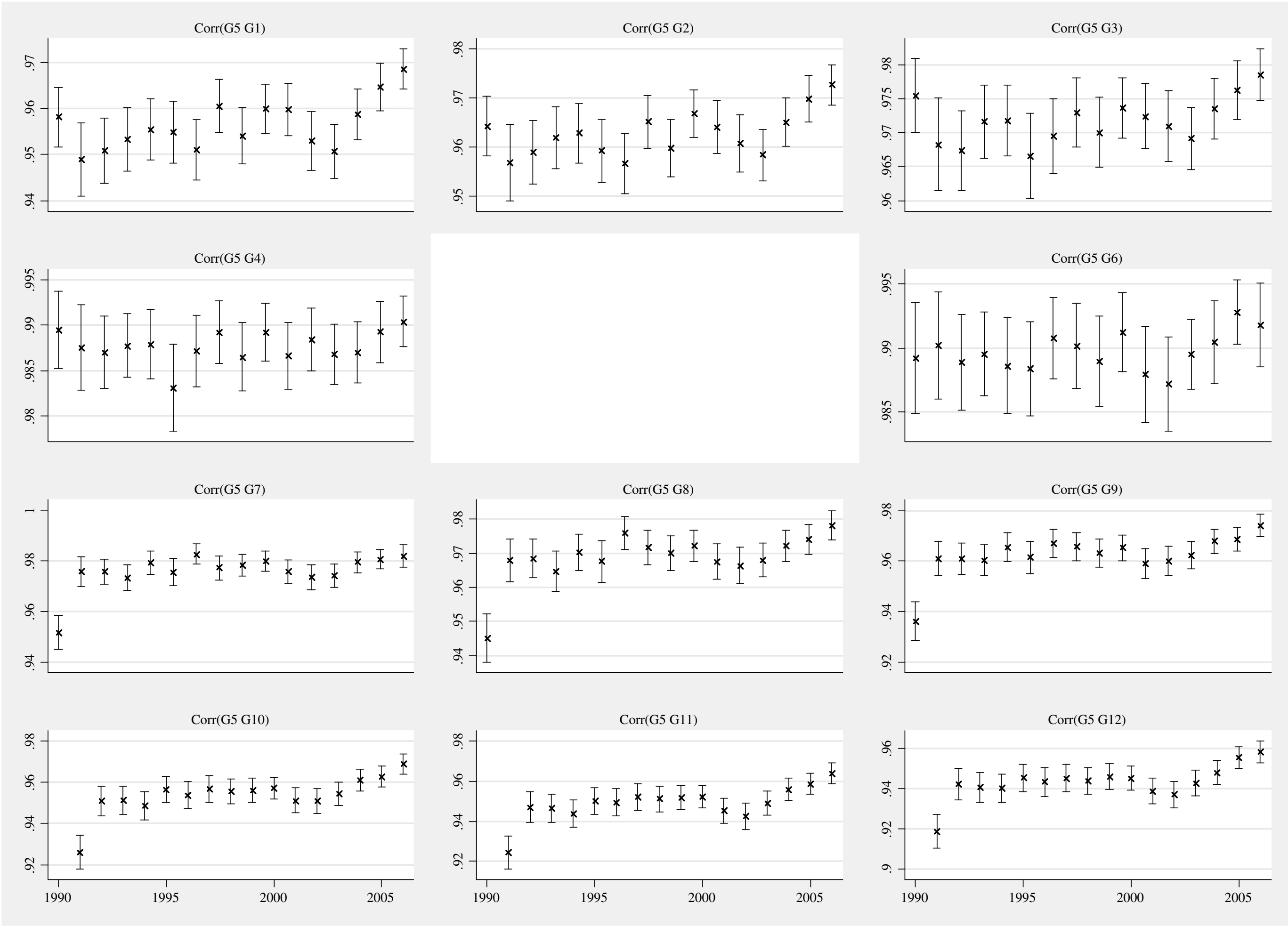
Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; “x”: point estimate.

Figure A2. Gini correlations for March with other months of the year



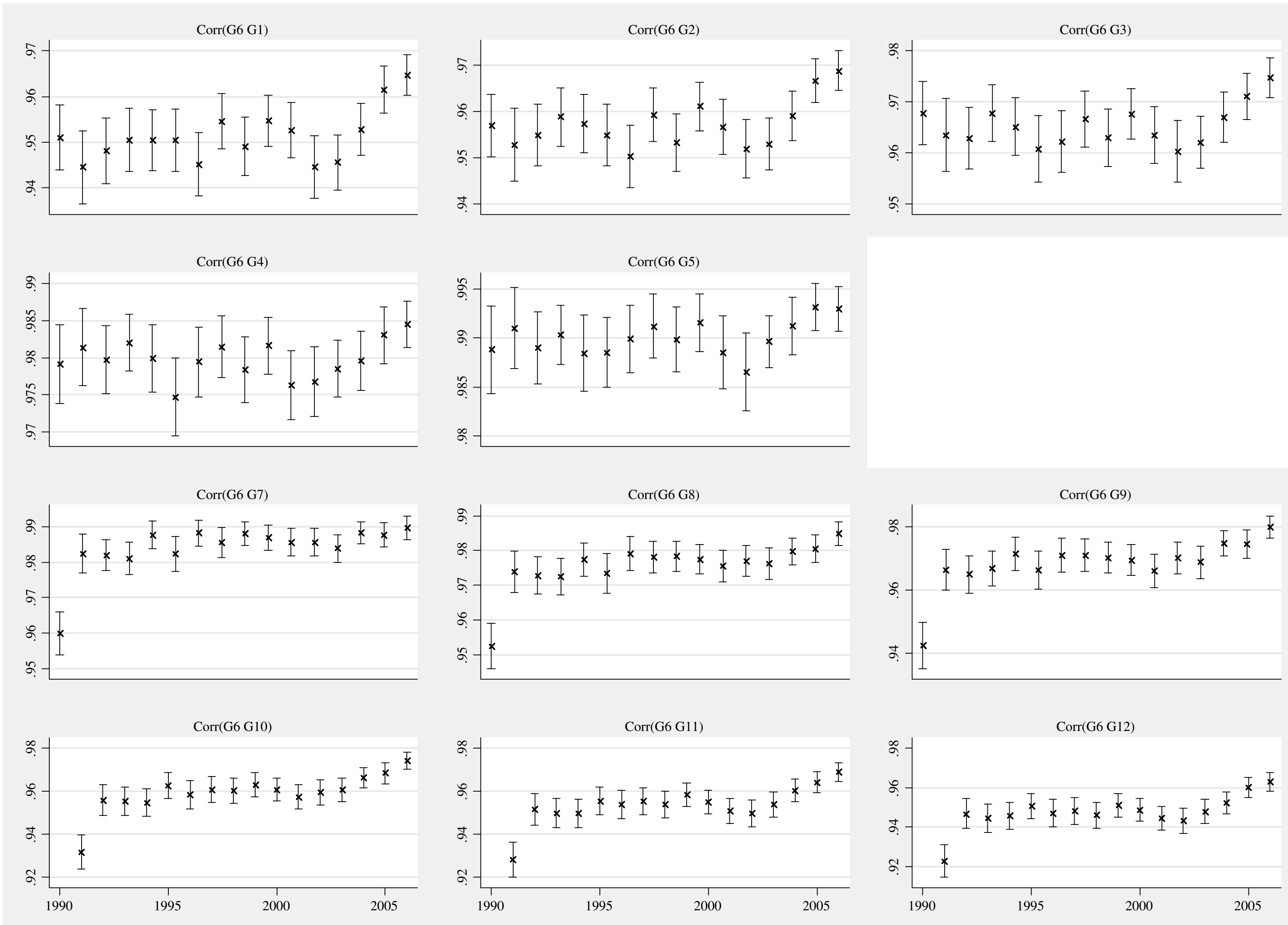
Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; "x": point estimate.

Figure A3. Gini correlations for April with other months of the year



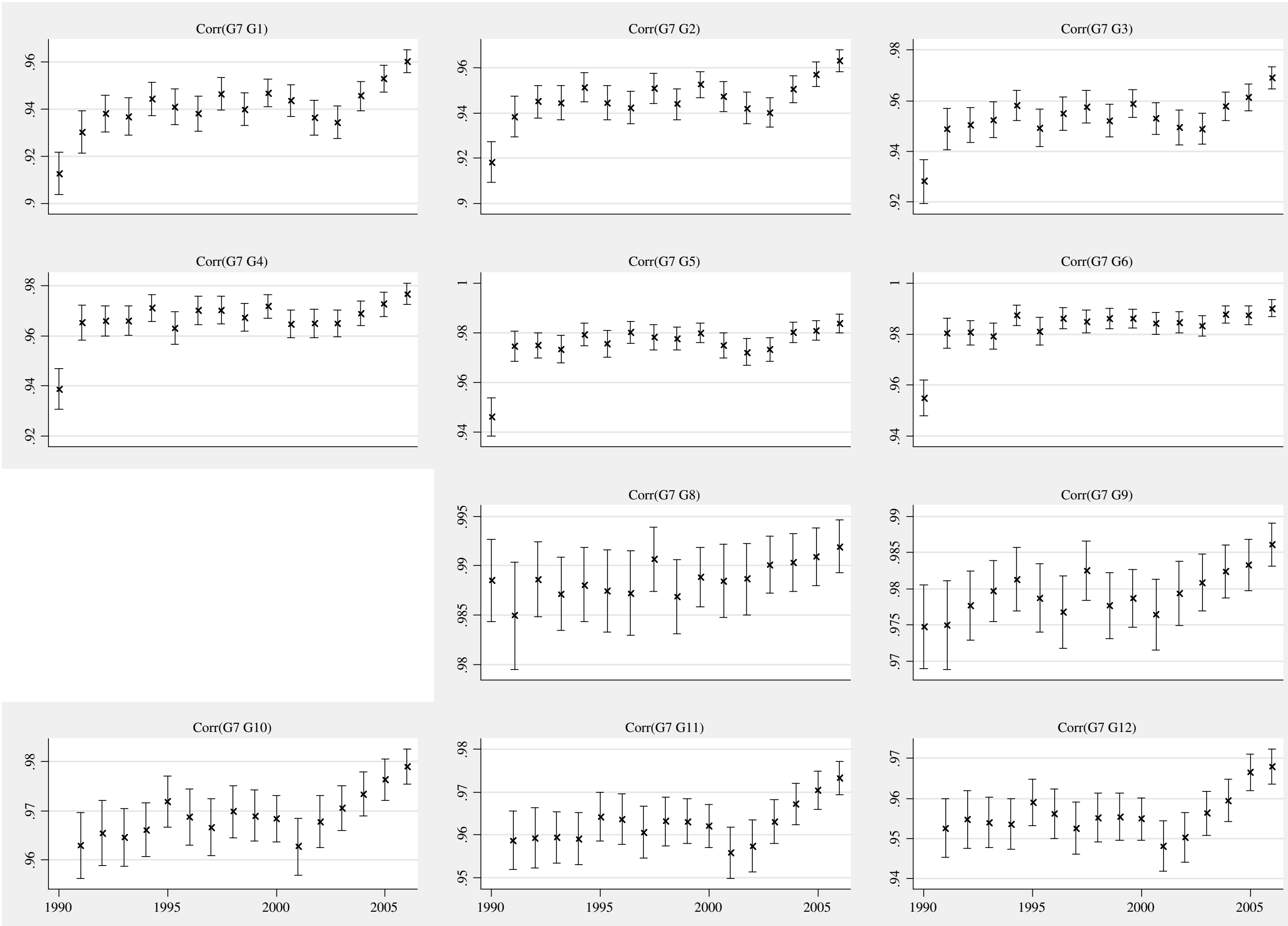
Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; “x”: point estimate.

Figure A4. Gini correlations for May with other months of the year



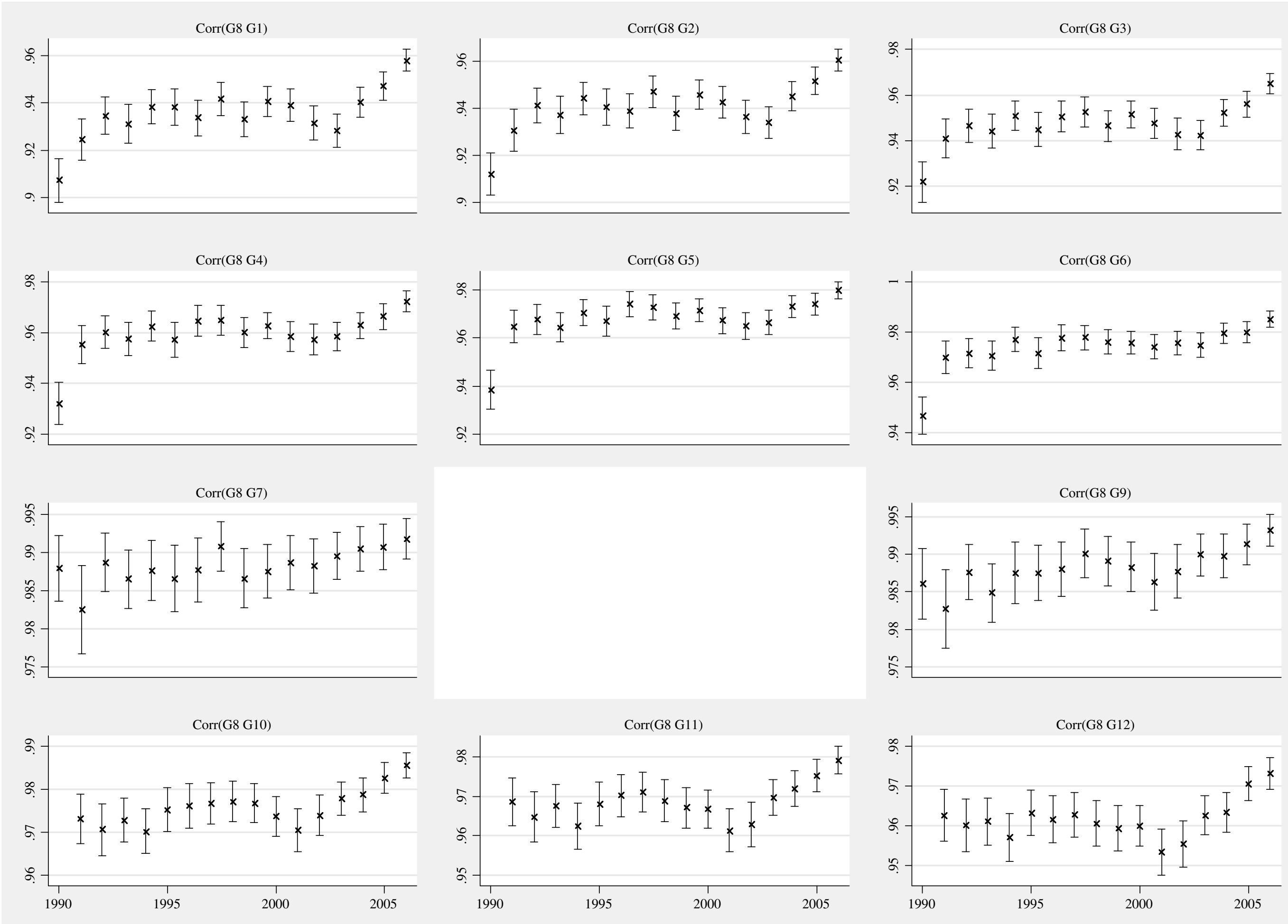
Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; “x”: point estimate.

Figure A5. Gini correlations for June with other months of the year



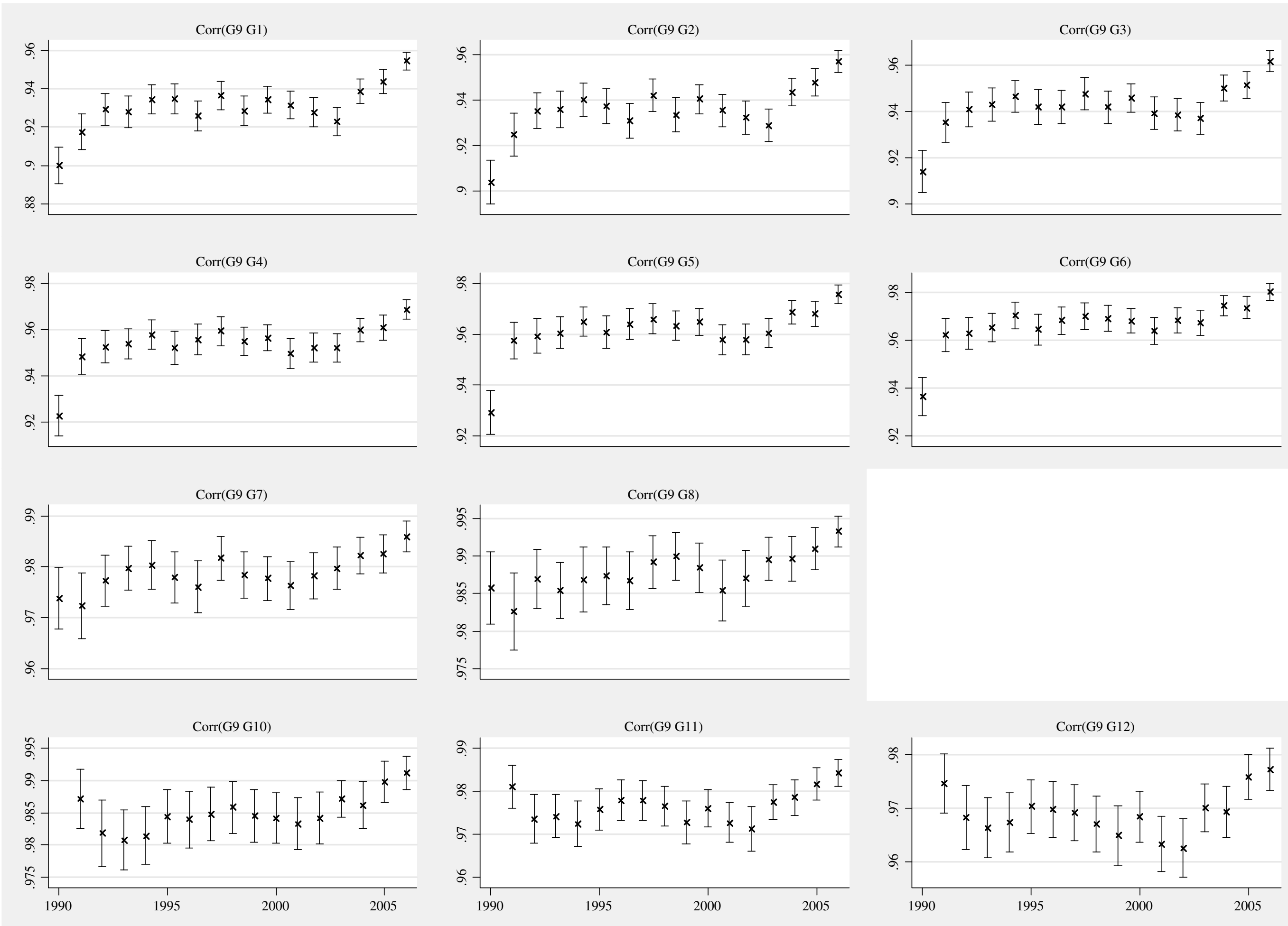
Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; “x”: point estimate.

Figure A6. Gini correlations for July with other months of the year



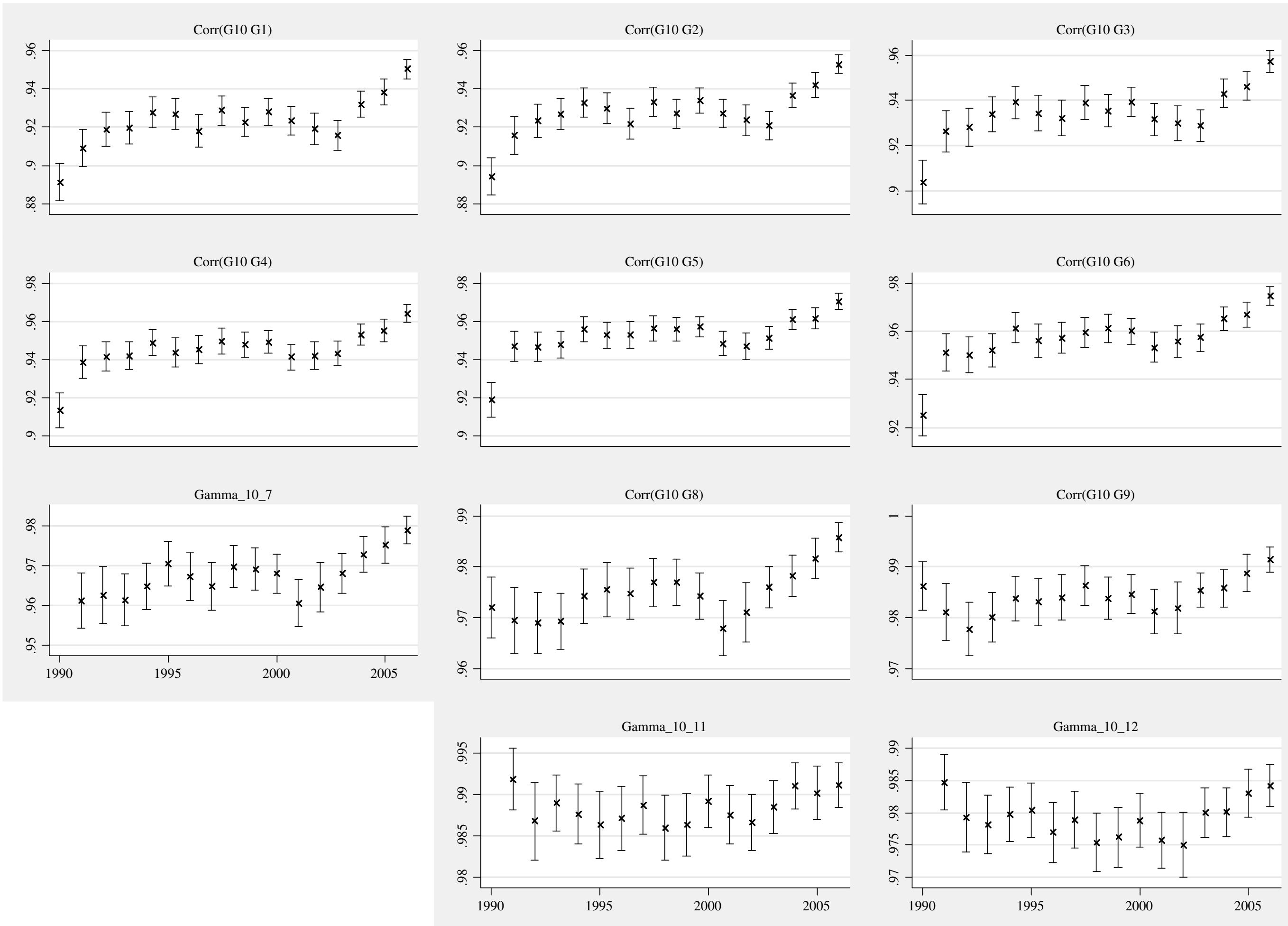
Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; "x": point estimate.

Figure A7. Gini correlations for August with other months of the year



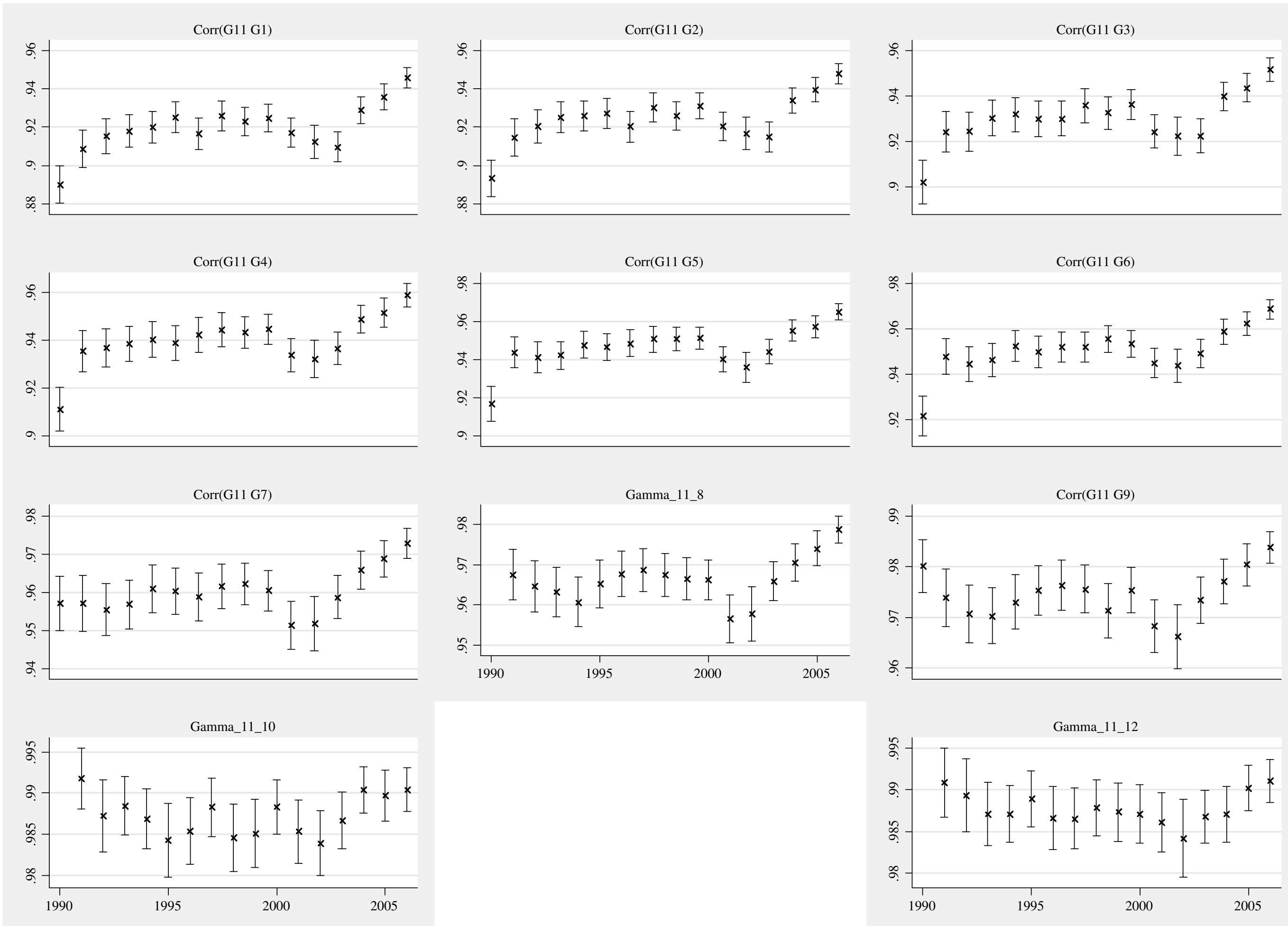
Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; “x”: point estimate.

Figure A8. Gini correlations for September with other months of the year



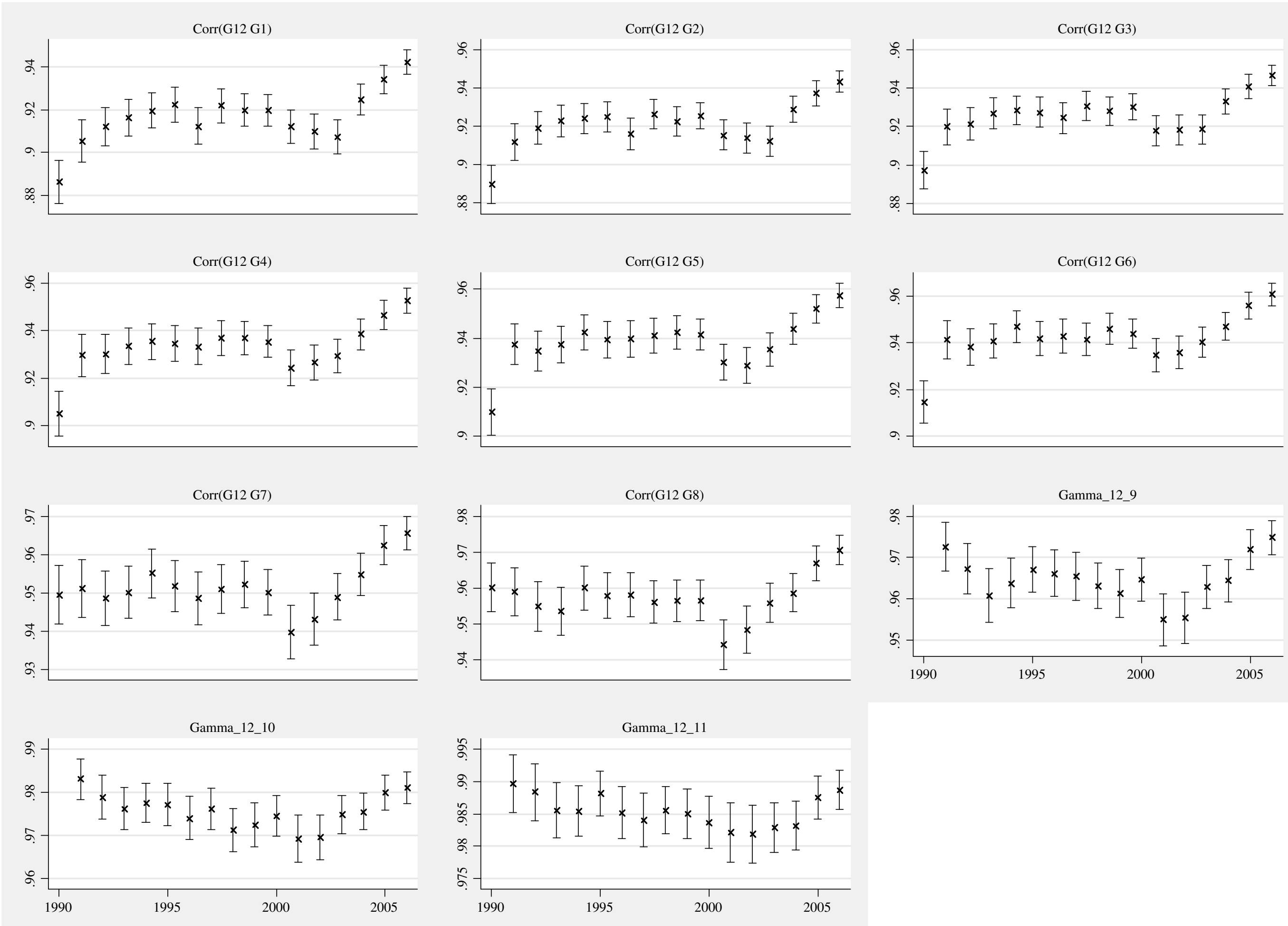
Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; “x”: point estimate.

Figure A9. Gini correlations for October with other months of the year



Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; “x”: point estimate.

Figure A10. Gini correlations for November with other months of the year



Note. Database is IAS 2006. Own calculations. Vertical bars: Jackknife confidence intervals; “x”: point estimate.

Figure A11. Gini correlations for December with other months of the year