It's not here – Look elsewhere! Macro panel data analysis of labour taxation and unemployment

Michael Feil
Institute for Employment Research (IAB),
Regensburger Str. 104, 90478 Nürnberg, Germany

April 25, 2012

Abstract

Do labour taxes cause unemployment? The existing macroeconomic evidence based on panel data for OECD economies generally seems to answer this question in the affirmative. There are, however, concerns about the robustness of the dominating modelling approach in the literature. Based on a rather systematic strategy and longer time series than previous studies, we do not find support for a positive effect of labour taxes on unemployment. This failure is in line with Granger causality tests also carried out. The degree of heterogeneity found for the countries in the sample raises serious doubts about the appropriateness of homogeneous panel models. The simple message resulting from this study is that the wanted (causal) link between labour taxes and unemployment is not where we looked for it.

Keywords: unemployment, labour taxation, panel econometrics

JEL Code: E24, H30, C33

Acknowledgements: I thank Oliver Krämer for excellent research assistance. Abbad El-Rayyes helped me in preparing the tables in LaTeX.

Contents

1 Introduction

Ever since the OECD Job Study (1994) taxes have been among the most suspected policy interventions that may cause unemployment. The theoretical arguments behind this hypothesis are manifold and beyond the scope of this paper. One rather general explanation can be found in the wage-push effect of income taxes in imperfectly competitive labour markets as described prominently by Layard, Nickell, and Jackman (1991). There is a lot of empirical evidence in support of positive effects of taxes on unemployment. See the accompanying paper (Feil, 2012) for an overview. A very convincing form of evidence would be given, if one could find a direct link between taxes and unemployment at the aggregate level. This form of macroeconometric evidence has to be robust in the sense that it does not depend on a few 'nice' specifications, however derived, but emerging from a rather systematic and general modelling strategy. Our reading of the literature is that such a systematic approach is missing or at least in scarce supply.

The aim of this paper is to answer a simple question: Do labour taxes cause unemployment?[1] The setting for the analysis could surely be manifold.[2] One could approach this question directly or indirectly, by linking taxes to unemployment, employment or wages. One could estimate systems of equations or reduced form models. One could use micro as well as macro data.

If one is taking the issue rather bluntly, a simple way of setting out the econometric analysis is to look for variation across time and space. Since 'unemployment' is naturally a macroeconomic concept, one might venture to take the topic very literally by linking labour taxes and unemployment as directly and as macro as possible. This is the plan we pursue here.

Obviously, this kind of "macro approach" is not new. Early contributions, similar to this one, can be traced back to the work of Bean et al. (1986) at least. The idea of using cross-country time series data in the field of (macro) labour economics gained momentum in the second half of the 1990s, when effort poured into the construction of long time series for many countries resulting in panels of about 600 observations (NT) for yearly data (e.g. Nickell et al., 2003). Taxes were one of multiple items in the regression analysis that became known as 'labour market institions (LMI)' (Nickell and Layard, 1999).

The literature that emerged on the basis of the institutional data (e.g. Nickell, 1997; Elmeskov et al., 1998; Blanchard and Wolfers, 2000; Nickell et al., 2003, 2005; Bassanini and Duval, 2009) has been mainly concerned with explaining unemployment, as well as − although with less intensity − employment and wages by LMI on the whole. Taxes seldom received special interest. However, when it comes to the findings, labour taxes rank rather high among the LMI that were found to have contributed significantly to the rise in unemployment. This holds with very few exceptions for the contributions mentioned above as well as for other studies (e.g. Belot and van Ours, 2001).

The conclusion reached by Nickell and Layard (1999) in their survey of *Labor Market Institutions and Economic Performance* that "[...]there is evidence that overall labor tax rates do influence labor costs in the long run and hence raise unemployment", is thus supported by

[2] For a survey of the literature in this respect see Feil (2012).

numerous studies. There are, however, critical contributions that doubt the robustness of the macro-panel approach as such (Howell et al., 2006; Baccaro and Rei, 2007).

One finding of the theoretical literature on taxation and unemployment, that has not been explored extensively yet, is the interaction of labour taxes and wage setting. The tax literature has identified different regimes that impact significantly, at least theoretically, on the effects of taxes on wages and thus (un-)employment. The short form of this argument is that income taxes and social security contributions are borne by workers in countries with flexible (competitive) labour markets while real-wage resistance is much more likely to occur in inflexible (uncompetitive) labour markets. For this reason the OECD (1994: 240) argued that "there is no particular reason for the effects of taxation on wage-setting to be equally significant in all countries". Consequently the OECD based much of its argumentation on the study by Tyrväinen (1995), who estimated vector autoregressive models for 10 countries separately.

Our reading of the literature is that there is a gap between the empirical macro LMI strand, assuming mostly homogeneous cross-country panel models, and the public economics strand, pointing towards significant differences between certain types of labour markets.

An obvious way to bridge the two strands would be to test the homogeneity assumption. Put differently one could give taxes in cross-country panel studies even greater explanatory power, if the total sample could be split into similar subgroups. The obvious drawback of such a 'heterogeneous approach' is the reduction in sample size. This trade-off is very likely responsible for the reluctance of the macroeconometric LMI literature to follow this idea. This holds in particular if researchers opted to collapse the data into five-year averages, an approach motivated by the low variation in some of the LMI variables (see e.g. Blanchard and Wolfers, 2000; Belot and van Ours, 2001).

With the available data now exhibiting a time dimension (T) of about 50, the case to pool the observations no longer is that strong.[3] Single-country models can now be estimated, even for a considerable number of explanatory variables. There are other methods in between like 'heterogeneous panels', to use the terminology of Baltagi (2008). So restricting the econometric analysis to homogeneous panel models is neither necessary nor a priori convincing.

The continuously growing T has another effect which has not been fully accounted for in the LMI literature. While Nickell's (1997) panel models with $T = 2$ in methodological respect belonged to the 'micro world', to borrow Eberhardt's (2009) terminology, a full-scaled panel model nowadays clearly belongs to the 'macro world'. This distinction is important since these are actually different methodological worlds with potentially different implications for the results.

The longer time series also bring up the risk of ill-specified econometric models due to nonstationary variables. This issue has been largely ignored in the previous literature with a few notable exceptions (Baccaro and Rei, 2007; Everaert and Heylen, 2002). Nickell et al. (2003) for instance add a note to their main results (p.14) stating that a unit-root test on the residuals rejects the hypothesis that these are integrated of order one in all countries. They take this as

[3]This argument obviously depends on the absence of structural breaks. If there are fundamental changes in the data generating processes then longer time series are of little value.

evidence for cointegration. From a systematic point of view (Campos et al., 2005, e.g.) this is not a completely convincing treatment of time series that were found to be nonstationary (Nickell et al., 2003, :214)[4]. Combining possibly integrated ratios, categorial variables, and measures of shocks, some of them stationary by construction, into one model appears at least problematic. A more thorough and systematic treatment could be deemed necessary.

A systematic approach to our topic is the second major contribution of this study, besides the application of newer estimation methods. While most of the existing studies have come up with a few nice estimation results, we try to look at the case of labour taxes and unemployment from a more systematic perspective.

One major aspect of this approach is a rather extensive analysis of (reverse) causality. Although a lot of economists might have fretted about the direction of causality there are only a few contributions that do actually try to test it.

A third, although not fully exploited contribution of the paper is an alternative measure of the income tax burden. Based on a representative-agent approach instead of macroeconomic aggregates we gathered data on average and marginal tax rates for a single-breadwinner couple with two children from 1972 to 2009.

Our results point to a vast amount of heterogeneity in the sample. The data for some countries support the hypothesis that higher taxes lead to higher unemployment. There is also evidence for no effects. Surprisingly the estimation results also allow for the possibility that higher taxes are a predictor for lower unemployment rates in some countries. With our data the perceived wisdom that overall labour tax rates raise unemployment can only be confirmed for a few special methodical cases, in particular by estimating homogeneous panel models with serially correlated errors.

The indeterminancy of the influence of taxes on unemployment in a multivariate model is reflected by the bivariate Granger causality tests that we perform. These tests do not provide much insight into neither the existence nor the direction of (Granger) causality between labour taxes and unemployment.

Altogether our findings should, however, not be taken as evidence against the hypothesis that labour taxes cause unemployment as advocated e.g. by Daveri and Tabellini (2000) or Prescott (2004). Our message simply is that it is hard to find evidence for a detrimental effect of labour taxes in cross-country time series data.

2 A first look at the data

At the beginning of the analysis a simple descriptive look at the data seems appropriate. While measures of labour market performance like employment and unemployment rates are rather obvious the measurement of taxation merits some justifications.

From tax theory most measures of taxation should be based on individual data. This claim

[4]Berger and Everaert (2009) analyse the data of Nickell et al. in detail. They find that the supposed cointegration "does not survive when small sample properties and heterogeneous cross-sectional dependencies are taken into account".

is simply due to the fact that it is the tax burden of (representative) workers and firms that influence the decisions of these agents. If there are rich micro data sets complete distributions of tax variables could be derived. Such an approach is, however, beyond the scope of this paper. Here we want to exploit the cross-country and time variation of taxes and labour market outcomes. This calls for a representative-agent approach or for summary measures of tax variable distributions.

Following the OECD's *Tax/Benefit Position of Workers* approach[5] we constructed time series for an average (production) worker. We take a single-earner employee, married, with two children, as our representative agent. For this household type it is possible to derive a time series for the period 1972 to 2009. The average income tax rate for the 20 OECD economies of the sample is presented in Figure 3.

The most important message of these charts is that there is no clear picture, no distinctive trend appears discernible. Exceptions to this general judgement are Australia, Austria and Belgium, which show a rising tax burden, while Germany, Ireland, Sweden, Switzerland and the US display a falling average tax rate. Some of the series show awkward spikes and declines. The overall mean is about 13 % with a standard deviation of almost 9 %. The highest rates are levied in the Nordic countries, the lowest in France, Japan and Portugal.

The observed marginal tax rates for the average worker encompasses an even larger interval than the average tax rates (Figure 4). However, its overall coefficient of variation is smaller. Again, there is no clear-cut pattern distinguishable from the complete sample. Some countries show a downward-sloping trend, like Sweden and the UK, while in others like Belgium and Germany marginal rates, at least for the particular unit under analysis, have risen.

Combining the average $\left(\frac{T}{Y}\right)$ and the marginal tax rates (T') in the coefficient of residual income progression (CRIP) defined as

$$CRIP = \frac{1 - T'}{1 - \frac{T}{Y}}$$

we obtain a measure for tax progression, which can be used to detect stylised facts. This exercise, however, leaves us also with a rather inconclusive picture (Figure 8). Due to the jumps in the marginal tax rates, which are caused by movements between tax brackets, the CRIP also fluctuates quite strongly in some countries. Nevertheless for Norway, Sweden and the UK a clear increase in the CRIP, i.e. a decline in tax progression, can be observed. At the other end of the spectrum Austria, Germany and Italy depict a more progressive tax system nowadays than in the 1970s.

Against the backdrop of the presented evidence on tax rates it comes without surprise that bivariate correlations between average (marginal) tax rates and unemployment rates are rather weak and totally inconclusive. Not surprisingly some of the panels in Figure 5 even display the counterintuitive result of a negative correlation between the average income tax rate and unemployment. The same holds for corresponding plots of the CRIP where we expect a positive

[5] The latest publication in the series is OECD (2011). The whole series dates back to the early 1970s, probably starting with OECD (1978)

correlation from wage-bargaining, efficiency-wage and search models, but also find examples of negative correlations, above all for Canada and Germany.

The evidence presented so far raises serious doubts about the appropriateness of our representative-agent approach to the measurement of taxation. Further concerns emerge when we look at the evolution of aggregate measures of the tax wedge like the ratio of labour taxes and employees' total compensation. These tax variables have been used extensively in the literature. In Figure 6 we plot the tax wedge (TW) as defined by Nickell et al. (2005) in the following way:

$$TW = \tau_{PR} + \tau_I + \tau_C$$

The derivation of the average payroll tax rate τ_{PR}, the average income tax rate τ_I, and the consumption tax rate τ_C is described in the appendix.

With the exception of Canada and the Netherlands there is an upward-sloping trend for all countries (Figure 6). Since the tax wedge is available from 1960 onwards we can enlarge the time dimension of the sample substantially and in this case a positive trend is found for Canada and the Netherlands, too.

We proceed by plotting the tax wedge against unemployment rates (Figure 7). In stark contrast to the scatter plots for the average tax rate (Figure 5) similar patterns emerge for all countries. In some economies the positive correlation is weak (e.g. for the US), but in general higher tax wedges involve higher unemployment rates.

There are signs of nonlinearity in the panels of Figure 7. In many countries there is flat part at the bottom left, indicating no influence of taxes on unemployment at low levels of taxation. With higher values of the tax wedge higher unemployment rates are associated, giving rise − at least in some countries like Austria, Belgium and Spain − to clearly visible nonlinearities.

Social security contributions (SSC), which are predominantly collected as direct levies on payrolls, should not be left out from the picture. Whether SSC are taxes or to what extent is discussed in the literature (Gruber, 1997). Their character obviously depends on the design of the social insurance benefits. If benefits are defined independently of individual contributions, which is typically the case in public health care, then SSC should be considered a tax. On the other hand if there is a strong link between contributions and benefits, i.e. fiscal equivalence, then they are rather insurance premia and should not be considered as taxes.[6]

Being aware of these limitations we note that SSC have increased in most countries of our sample (Figure 9). The rates depicted are derived in the same way as the average and marginal tax rates before, i.e. from individual data. However, since SSC are most often proportional levies they provide evidence beyond the average worker. Again, the comparative analysis reveals large differences. While in New Zealand there is not any levy, continental European countries feature rates above 20%. Since the beginning of the series, SSC rates have increased in most OECD countries. This does not hold for Ireland, Norway, and the Netherlands which show

[6]More precisely the nature of SSC depends on individual characteristics and preferences and is thus very heterogeneous across workers. Uniform unemployment insurance (UI), for instance, acts like a tax for those workers who face only a small risk of becoming unemployed. At the other end some workers are heavily subsidised by UI since they would not have access to private insurance at typical UI rates.

significant declines in the rates in the second half of the period covered. The data also hint at upper ceilings towards the end of the series, with Germany and France as the most obvious examples. These 'upper limits' could be signs of deliberate policies to contain further increases, which have become suspect of causing (additional) unemployment, at least since the publication of the OECD's (1994) *Jobs Study.*

The patterns of correlation between unemployment and SSC rates are similarly inconclusive as those between average income tax and unemployment. There are some exceptions though. The evidence for France, Germany, and surprisingly Ireland supports the hypothesis of a payroll tax driven rise of unemployment. If we combine SSC and average tax rates into one measure, Germany will drop from the list. However Austria, Italy, Japan, and the Netherlands can then be named as empirical support.

Finally we also look at changes in unemployment rates and changes in the tax wedge. Figure (10) is inspired by a similar chart in OECD (1994), but looks at the differences between 1960 and 2008 compared to 1978 to 1991. Over this long period of time there are some countries − Denmark, Finland, and Sweden − that seem to support the supposed positive correlation between (changes) in labour taxes and (changes) in unemployment. The US complete the picture in the lower left corner.

This kind of presentation of the data depends significantly on the specific period of time chosen. For instance, if we repeat the exercise for 1972 to 2008, including New Zealand, Norway, and Spain, for which data is available in 1972, the positive correlation depends on Finland, Spain, Sweden, and the Netherlands on the one hand, and Canada and New Zealand on the other hand. Spain could be considered an irrelevant outlier, due to its authoritarian regime back then. But the correlation persists for other periods of time.

The OECD back in 1994 concluded that neither approach, whether in levels or in differences "shows a significant correlation between taxation and unemployment"(p.244). One has to bear in mind though, that this judgement is based on two cross sections. Thus the OECD's conclusion merely is that a stable positive correlation between unemployment and labour cannot be detected by simply looking at the variation across OECD economies at one point in time. We thus take the analysis one step further and plot the first differences ($\Delta ur_t = ur_t - ur_{t-1}, \Delta tw_t = tw_t - tw_{t-1}$) by country. This exercise reveals (Figure 11) patterns of 'no correlation' for almost all countries of our sample.

Eventually, looking at all countries and years leaves us with a totally inconclusive bunch of data points around zero. The correlation coefficient for all pairs of all changes in unemployment and taxes is $-.05$. For the data in levels it is .28.

The bivariate evidence presented in this section should be taken as a first indication that the (causal) relationship between our main variables could be difficult to establish. This appears to be particularly warranted for the data in first differences. From the set of tax variables only the overall tax wedge in levels shows a positive correlation with unemployment. For this reason we are going to concentrate on the tax wedge in what follows.

3 Direction of causality

A positive correlation between taxes and unemployment provokes arguments about the direction of causality. This holds in particular for unemployment insurance rates, an element of our tax wedge variable, which typically rise with the level of unemployment unless the additional means needed in times of high joblessness are taken from other sources of public revenue.

Formal tests of causality between taxes and unemployment are rare in the literature. An early example is Elmeskov et al. (1998) who report tests for reverse causality in the appendix. There are three countries (Austria, Ireland, and Norway) where the data support the hypothesis that the tax wedge is (Granger) caused by unemployment. Recently, Rault and Vaubourg (2011) systematically tested for (Granger) causality of unemployment and labour market institutions in connection with financial institutions. The only significant causal link they find is for Ireland, where a higher tax wedge causes unemployment.

The standard approach to test for causality is based on Granger (1969). According to Granger's definition a time series X_t causes another time series Y_t if in a regression of Y_t on X_t as well as past values of X_t, i.e. X_{t-l}, reduces the forecast error variance of Y_t significantly.

In practice Granger causality tests are carried out by estimating a vector autoregressive (VAR) model of the form:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + ... + \alpha_p Y_{t-p} + b_1 X_{t-1} + ... + b_p X_{t-p} + u_t \quad (1)$$

$$X_t = c_0 + c_1 X_{t-1} + ... + c_p X_{t-p} + d_1 Y_{t-1} + ... + d_p Y_{t-p} + v_t \quad (2)$$

Then, testing $H_0 : b_1 = b_2 = ... = b_p = 0$, against $H_1 : b_l \neq 0 \; \exists l \in \{1, ..., p\}$, is a test that X_t does not Granger-cause Y_t. Additionally, testing $H_0 : d_1 = d_2 = ... = d_p = 0$, against $H_1 : d_l \neq 0 \; \exists l \in \{1, ..., p\}$, is a test that Y_t does not Granger-cause X_t.

With our data set we could either test for causality for each country separately or we can resort to panel Granger-causality (GC) testing as e.g. in Hartwig (2010). Due to the heterogeneity found in the analysis of correlation above we consider both.

Finally, it is important to note that the causality tests applied in the following assume that all other information is irrelevant, whether it is available or unavailable to the observer. This assumption is strong in our case since there are sound theoretical reasons to expect wage-setting institutions to influence the effect of taxes on unemployment.

3.1 Granger-causality tests for each country

To test for causality we run four sets of regressions. In all of them standardised OECD unemployment rates (ur) and the tax wedge (tw) introduced above are used. The four approaches differ in the way they take account of the (potential) order of integration of the two series. Whether or not rates could be I(1) is sometimes disputed. We take an agnostic stance by simply running alternative specifications, one in levels, one in first differences and another one based on standard unit root tests. In the later the possibility of a mix of one series in levels and the other in first differences arises.

The fourth set of tests follows a different approach (Toda and Yamamoto, 1995). The test VAR is always estimated in levels. However, the order of integration determines the number of additional lags included in the VAR. That means that we add one more lag as otherwise indicated by the usual methods, if the series has been found to be I(1). We employ the Schwarz information criterion to determine the appropriate number of 'base lags'. In the coefficient tests only these 'base lags' are included.[7]

Employing augmented Dickey-Fuller tests for unit roots in the unemployment rates (Table 8), we find that the null hypothesis, i.e. the single country time series contain a unit root, can only be rejected for Germany, Sweden, Switzerland, and the US. Except for Switzerland this finding is confirmed by the KPSS test (Kwiatkowski et al., 1992) for stationarity (Table 10). However, the KPSS test indicates stationary unemployment rates for Austria, Finland, Ireland, and Portugal, too.

The properties of the tax wedge series cannot be determined precisely, too. Here, the KPSS tests suggest that all series are nonstationary (Table 9). The ADF tests do not indicate a unit root for all countries, though (see Table 11 in the appendix for details).

In all cases of contradictory results from the ADF and KPSS tests we make use of the Phillips-Perron test (Phillips and Perron, 1988). For the unemployment series this test confirms the results of the ADF tests. For the tax wedge the additional information provided suggest that the series for Belgium, Canada, Norway are I(0) while all the others are I(1).

	Directions of causality						
	Taxes $\rightarrow$ Unemployment		Unemployment $\rightarrow$ Taxes		Both		None
	No.	Names	No.	Names	No.	Names	No.
Levels	6	AUS, CAN, FIN, GER, NEL, SWZ	3	BEL, FRA, USA	2	DEN, NEW	9
1st Differences	3	CAN, NEL, UK	4	FRA, IRE, JAP, USA	1	AUS	12
Based on unit root tests	3	CAN, NEL, UK	3	FRA, IRE, JAP	2	AUS, USA	12
T-Y-approach	4	AUS, CAN, NEL, UK	5	BEL, FRA, IRE, ITA, NEW	2	DEN, USA	9

Table 1: Country-specific Granger-causality tests

The results of the four sets of country-specific Granger causality tests are summed up in Table 1. They do not differ strongly for the alternative treatment of the time series (rows). The important result is that for the majority of countries the tests fail to reject the null hypothesis for both models, thus nothing can be said about causality. There are some 'robust' cases where the GC tests agree. The case for causality running from taxes to unemployment is most clearly supported by the evidence for Canada and the Netherlands. The case for 'reverse causality' is made by France. Our preferred set of tests – T-Y-approach – finds a relative majority for the hypothesis that unemployment causes higher taxes.

There are no signs of 'suspected clusters' in the results like Anglo-Saxon or Continental-Europe.

[7]That means that we cannot use Stata's *vargranger* command. Hence we employ the standard *test* function.

3.2 Panel Granger-causality tests

Our panel data allows for the possibility of causality tests to be based on many more observations. However, depending on panel heterogeneity this could be an advantage or disadvantage. Taking related econometric problems seriously (see Pesaran and Smith, 1995; Weinhold, 1999) we follow two different paths. At first we estimate fixed-effects dynamic panel models and test for Granger-causality in the standard way. We then apply the method of Weinhold (1999) and Nair-Reichert and Weinhold (2001) and estimate mixed fixed- and random models which have some advantages if the panel is heterogeneous. We are going to take up the issue of panel heterogeneity explicitly later.

Both methods used assume stationary variables. We thus apply standard panel unit root tests first (see Table 12 in the appendix). The results of the Fisher-type ADF- and Phillips-Perron tests support the null that there are unit roots in all unemployment series. This conclusion seems questionable when the Levin-Lin-Chu test is taken into account. However, since in case of cross-country dependence there could be size-distortions and lack of power of the standard unit root tests, which assume cross-sectionally uncorrelated panels, we also apply Pesaran's test (2007), also termed CIPS test (cross-sectionally augmented IPS). This unit root test does not contradict the null of a unit root in all panels. The cross-check provided by the Hadri test confirms this finding. We thus conclude that although there might be exceptions for some countries, the unemployment rates appear to be I(1).

For the tax wedge the null hypothesis that there are unit roots in all series appears to be too restrictive (Table 12). On the other hand the Hadri test for stationarity indicates that some panels include unit roots. These findings are largely in line with the results from the country-specific unit root tests. We conclude that in order to assure stationarity both series have to be differenced once.

The first set of GC tests is based on the following model:

$$\Delta ur_{it} = a_0 + \sum_{l=1}^{p} a_l \Delta ur_{it-l} + \sum_{l=1}^{p} b_l \Delta tw_{it-l} + \mu_i + e_{it} \tag{3}$$

and the corresponding formulation for the opposite direction of causality. (3) is a fixed-effects panel model with country-specific means μ_i. The error term e_{it} has the usual iid-properties with zero mean. The lag length, i.e. the order of p is determined by the Schwarz Information Criterion (SIC).

As is well known in the literature there is an endogeneity problem in dynamic panel models. Hence (3) cannot be estimated consistently by OLS. Solutions to the endogeneity problem are provided by the Arellano-Bond estimator (see Arellano and Bond, 1991) and the Arellano-Bover/Blundell-Bond estimator (see Arellano and Bover, 1995; Blundell and Bond, 1998). Note, however, that these estimators were developed for panels with a large N and a small T. With a growing time dimension the bias diminishes, so we also estimate the models by OLS as an additional check.

Table 2 provides a summary of the panel Granger causality tests based on different estimation methods. There are no signs of causality as running from labour taxes to unemployment. The same holds true for tests for the opposite direction of causality (Table 3).

	(OLS) Δur	(Arellano-Bond) Δur	(AB/BB) Δur
Δur_{t-1}	0.550***	0.548***	0.563***
	(8.26)	(8.44)	(3.89)
Δur_{t-2}	−0.199***	−0.200***	−0.165
	(−3.89)	(−4.02)	(−0.81)
Δtw_{t-1}	1.435	1.508	3.591
	(0.67)	(0.71)	(0.11)
Δtw_{t-2}	1.761	1.794	3.895
	(1.37)	(1.45)	(0.32)
Constant	0.0677***	0.0675***	0.0395
	(5.31)	(4.88)	(0.22)
Observations	947	927	947
Wald_p	0.000	0.000	0.000

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2: Panel Granger-causality tests: unemployment rates

The non-existence of evidence for causality in either way is confirmed by the test strategy suggested by Hsiao et al. (1989) and Weinhold (1999). The method is applied in a comparable setting by Nair-Reichert and Weinhold (2001), who propose to use the following 'Mixed Fixed and Random (MFR)' model

$$\Delta ur_{it} = a_0 + \gamma_i \Delta ur_{it-l} + \beta_{1i} \Delta tw^{o}_{it-1} + \mu_i + e_{it} \tag{4}$$

where μ_i is a fixed effect and $\beta_{1i} = \bar{\beta}_1 + \eta_i$ is a random coefficient with $\eta_i \sim N\left(0, \sigma^2_{\eta}\right)$. The coefficient of the lagged-dependent variable is country-specific and fixed.

The test is based on a conventional Wald test of the 'candidate causal variable' as Δtw^{o}_{it} is termed by Nair-Reichert and Weinhold.[8] The superscript o denotes the fact that Δtw_{it} is orthogonalized before it enters (4). It is constructed by regressing Δtw_{it} (X_{it} in general) on other right-hand side variables and lags of Δtw_{it} (X_{it}). According to Nair-Reichert and Weinhold "orthogonalization is necessary to ensure that the coefficients are independent ..." (p.8). The approach can also handle further covariates, but we restrict the model to unemployment and taxes. Applying the technique we neither find evidence for causality of taxes on unemployment nor the other way round (Table 4).

Nair-Reichert and Weinhold argue that their approach is a diagnostic tool to detect the magnitude of panel heterogeneity, because the size of the estimated variance relative to the

[8]In general notation we would use X^{o}_{it}, indicating the supposed (weak) exogeneity or pre-determination of the supposed causal variable.

	(OLS) Δtw	(Arellano-Bond) Δtw	(AB/BB) Δtw
Δur_{t-1}	-0.0000910	-0.000113	-0.000320
	(-0.08)	(-0.11)	(-0.28)
Δur_{t-2}	0.000470	0.000563	0.00115
	(0.61)	(0.75)	(1.39)
Δtw_{t-1}	0.0205	0.0116	-0.155
	(0.37)	(0.21)	(-0.53)
Δtw_{t-2}	-0.0249	-0.0312	-0.00265
	(-0.63)	(-0.75)	(-0.03)
Constant	0.00451^{***}	0.00457^{***}	0.00475^{**}
	(14.11)	(6.07)	(3.02)
Observations	947	927	947
Wald_p	0.729	0.600	0.373

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3: Panel Granger-causality tests: tax wedge

Variable	coefficient	std. error	σ_η	estimated prob. of causality
dependent Δur_{it}				
Δtw^o_{it-1}	1.73	3.02	9.88	0.57
dependent Δtw_{it}				
Δur^o_{it-1}	-.00015	.00067	.0029	0.64

Table 4: Panel GC tests − Nair-Reichert/Weinhold approach

size of the (random) coefficient is informative in this respect. Comparing the estimate of the variance of the random coefficients, or the standard deviation σ_η, we note that in both models this figure is (much) higher than the estimated mean coefficient $(\bar{\beta}_1)$.

This evidence is in line with the single-country Granger causality test of the previous section where we could not find a clear pattern for the whole sample.

4 Time-series-cross-section analysis

A striking common feature of empirical panel studies on the influence of labour taxes − or more generally labour market institutions − on unemployment is the homogeneity assumption (see e.g. Blanchard and Wolfers (2000); Nickell et al. (2005)). There are some examples of a sample split and estimation for separate sub-groups (see e.g. Daveri and Tabellini (2000)) and also a few cases of country-specific estimates (Tyrväinen, 1995; García and Sala, 2008). Restricting the coefficients of a single LMI across time and space to a single number seems a priori very limiting. So one could expect some justification for the popular homogeneous panel approach. The literature, however, is rather silent on this issue. Mostly there is not more than an explicit tests for poolability, as e.g. in Nickell et al. (2005).

We are going to take up this issue in depth in the section after the next. Before that we first take a closer look at the specification of the model to be estimated.

4.1 Derivation of the estimated equation

The empirical literature is dominated by models of the generic form

$$ur_{it} = \alpha_i + \boldsymbol{Z}'_{it}\boldsymbol{\beta} + \varepsilon_{it} \tag{5}$$

where $\boldsymbol{Z}_{it}$ is a vector of wage-pressure variables, including the replacement rate, (trade union) bargaining power, and labour taxes among others. One striking property of (5) is the lack of any wage variable. This can be justified by the following simple model (see Nickell et al., 2003, p. 415). If the wage-setting equation has the form

$$w - p = \alpha_0 - \alpha_1 ur + \alpha_2 z \tag{6}$$

and the labour-demand (or price-setting) equation has the form

$$p - w = \beta_0 - \beta_1 ur \tag{7}$$

where w denotes wages, p prices, and $\{\alpha_k\}$ and $\{\beta_j\}$ are two sets of parameters, then equilibrium unemployment (ur^*) is given by

$$ur^* = \frac{\alpha_0 + \beta_0 + \alpha_2 z}{\beta_1 + \alpha_1} \tag{8}$$

The notion of equilibrium unemployment is crucial here. For an empirical application (5) would require strong assumptions on the speed of adjustment of wages, particularly if annual data are used.

Another version of (5), demanding less justification, is a simple dynamic extension of it, including the lagged dependent variable

$$ur_{it} = \alpha_i + \gamma ur_{it-1} + \boldsymbol{Z}'_{it}\boldsymbol{\beta} + \varepsilon_{it} \tag{9}$$

The alternative approach to such a reduced form model is a structural model, consisting of a labour-demand and a wage-setting equation. Estimation of such a structural model typically requires exclusion restrictions in order to identify both equations of the system (see Manning, 1993). Examples of this approach are e.g. Steiner (1998) and Boockmann et al. (2001).

The inclusion of the lagged unemployment rate is also justified by its potential to account for omitted variables. The most obvious kind of variables missing from (9) are measures of productivity. If the production function of the representative firm in the underlying wage-bargaining model is not Cobb-Douglas, then there are good reasons to augment the estimated equation by a productivity variable. The inverse of the capital output ratio, $\left(\frac{Y}{K}\right)$ or $(y - k)$ in logarithms, would be such an extension (see Manning, 1993, p. 106f.).

Another direction in which (9) can be developed is explored by the 'shocks and institutions' literature (see e.g. Blanchard and Wolfers, 2000; Nickell et al., 2005). A linear extension to (9) takes the form of

$$ur_{it} = \alpha_i + \gamma ur_{it-1} + \boldsymbol{Z}'_{it}\boldsymbol{\beta} + \boldsymbol{X}'_{it}\boldsymbol{\delta} + \varepsilon_{it} \tag{10}$$

where X summarises macroeconomic level variables like real interest rates or (mean-reverting) macro shocks typically derived as deviations from a long-term trend.[9] Finally we should note that (10) can also include interactions between institutions and macro variables (see e.g. Belot and van Ours, 2001).

While there is certainly a theoretic basis to include macroeconomic shock variables in an unemployment equation (see e.g. Nickell, 1998), most specifications used appear to be derived in an ad-hoc way. This interpretation of the literature is supported by some of the main contributors themselves, like Blanchard and Wolfers (2000), who acknowledge that "[t]he specification of [the estimated model] is clearly more a description of the data than the outcome of a tightly specified theory of interactions" (p. C19). And Nickell (1997, p. 65) in a more general comment on the nature of studies on cross-country correlations adds that "we see them [the estimated equations] as a helpful overview of the correlations in the data and nothing more".

Along this line of reasoning we stick to the class of linear models given by (10). The shocks-and-institutions specification obviously nests the 'institutions-only' model as a special case which we are going to look at, too.

[9]For a nonlinear specification see Blanchard and Wolfers (2000).

4.2 Data

Following the literature on the effects of labour market institutions (LMI) in a cross section of countries as discussed in the previous paragraph we now turn to the issue of data construction and model selection. Comparative data on LMI is scarce. The major source for time series is the OECD, which now provides a lot of information including indicators for employment protection, benefit replacement rates and active labour market policies.[10] Information on the institutional features of the wage-setting process are provided by the Amsterdam Institute for Advanced Labour Studies (AIAS) through their ICTWSS database.[11]

Since we want to use time series as long as possible we resort to the *Labour Market Institutions* database by Nickell and Nunziata (2001), which includes data from different sources, above all from the OECD, but also contributions from other research papers. Nickell and Nunziata's series start in 1960, although not all variables for all countries could be traced back until then. Detailed descriptions of the series can be found in the appendix.

Two LMI variables merit further remarks. The employment protection indicator (ep) as well as the wage-bargaining coordination (bc) measures are by definition categorical variables. The ep-indicator is constructed (from 1985 onwards) from a rather large information set. It is roughly continuously distributed in the interval [0;2]. The bc-indicator can take on five different values, ranging from 'fragmented bargaining' to 'economy-wide bargaining'. Both variables display a distinct ordering of their values and can thus be considered as ordinally scaled. While this pecularity has not received much attention in the literature, we remain skeptical as to whether both variables can be treated as being continuous, the standard approach in most applications. Instead of simply assuming continuity we convert ep and bc into dummy variables and analyse the estimated coefficients in the various model in more detail. This treatment also allows the bargaining coordination measure to display non-linear effects on unemployment as hypothesised by Calmfors and Drifill (1988).

The macroeconomic (shock) variables originate almost unanimously from the OECD. We consider the following variables: (i) a TFP shock, (ii) a labour demand shock, (iii) the output gap, (iv) a terms-of-trade shock, and (v) real interest rates. The construction of the shock variables is also described in the appendix.

From similar estimations (Eichhorst et al., 2010) we know that both sets of variables, shocks and institutions, are plagued by multicollinearity. Thus we also resort to principle component analysis (PCA) to reduce the dimensionality of our data.

4.3 'General unrestricted model'

The basic guideline in what follows is the 'general-to-specific' (*Gets*) approach as set forth by Hendry (1993) and described for instance in Campos et al. (2005). Thus we start with a linear model which is in principle only restricted by the available variables in our dataset. Obviously, whether to pool or not to pool is a critical issue in the model selection process. Our

[10] Large parts of the data are available on http://stats.oecd.org/.

[11] See http://www.uva-aias.net/207.

general unrestricted model (GUM) in the class of linear regression models, to which we confine the analysis, is a system of country-specific models allowing for correlation in the unobserved part. In this case no cross-equation restrictions are opposed, but the set of individual equations nevertheless gives rise to a system to be estimated jointly for increased efficiency, i.e. the seemingly unrelated regression (SUR) model by Zellner (1962).

4.3.1 Seemingly unrelated regression models

We start with the following model, which can be conceived as the SUR version of (10).

$$\begin{bmatrix} \Delta \boldsymbol{ur}_1 \\ \Delta \boldsymbol{ur}_2 \\ \vdots \\ \Delta \boldsymbol{ur}_N \end{bmatrix} = \begin{bmatrix} \Delta \boldsymbol{ur}_{1_{-1}} \\ \Delta \boldsymbol{ur}_{2_{-1}} \\ \vdots \\ \Delta \boldsymbol{ur}_{N_{-1}} \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_N \end{bmatrix} + \begin{bmatrix} \boldsymbol{Z}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{Z}_2 & \cdots & \mathbf{0} \\ & & \vdots & \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{Z}_N \end{bmatrix}' \begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \\ \vdots \\ \boldsymbol{\beta}_N \end{bmatrix} + \mathbf{X}'\boldsymbol{\delta} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_N \end{bmatrix} \quad (11)$$

Here $\boldsymbol{ur}_i$ denotes the vector of observed unemployment rates for country i. The observations of the K institutions are collected in $\boldsymbol{Z}_i$ matrices of dimension $(K \times T)$, those of the macroeconomic variables in the matrices $\boldsymbol{X}_i$ $(L \times T)$. The country-specific parameter vectors $\boldsymbol{\beta}_i$ and $\boldsymbol{\delta}_i$ are $(K \times 1)$ and $(L \times 1)$, the error term vector $\boldsymbol{\varepsilon}_i$ is of dimension $(T \times 1)$. The SUR model assumes $E[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'] = \mathbf{V} = \boldsymbol{\Sigma} \otimes \mathbf{I}$. The covariance matrix $\boldsymbol{\Sigma}$ is $(N \times N)$ with single element σ_{ij}.

Since the tests for (non-)stationary lead to the conclusion that besides unemployment rates and tax wedges, the series of trade union membership (*tud*) and benefit replacement rates (*brr*) are also − at least predominantly − integrated of order 1, we decided to diffence all continuous institution variables once.

Figure (1) summarises the estimated coefficients of the tax wedge variable for a subset of 16 countries. New Zealand, Portugal, Spain, and Switzerland had to be excluded due to missing data on total factor productivity (shocks). Apparently, the parameter estimates differ considerably between OECD countries. The maximum estimate of 11.2 (std. err. 5.13) is recorded for Belgium. The minimum of -18.9 (std. err. 5.97) is found for Germany. Most of the estimates are are not significant. None of the estimates within the interval $[-5, 5]$ are different from zero at conventional levels of significance.

The results of the SUR model are in line with the eyeballing-evidence of the section 2 and the findings of the Granger-causality tests. There seem to be enormous differences in the correlations between taxes and unemployment across countries. Surprisingly, the mean of the estimated tax wedge coefficients is negative.

The model in first differences shows a decent fit with R^2-values between .5 and .9 with the exception of Austria. However, the majority of the estimated coefficients is found to be insignificant in most countries. A rather obvious reason could be the low degree of freedom that results from $T = 42$ and $K + L \in \{10...14\}$ depending on the variation in the dummy variables for ep and bc. So one direction that should be explored to improve the SUR model is the reduction of the set of explanatory variables.

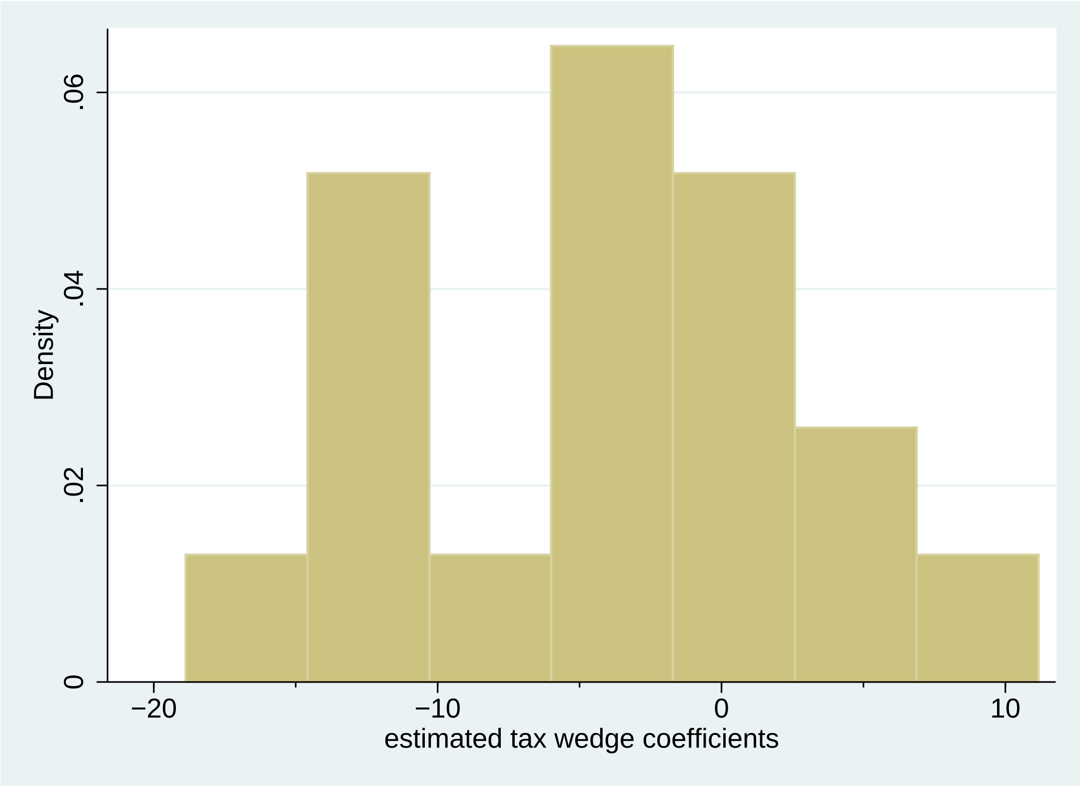

Figure 1: Distribution of SUR estimates in 'shocks-and-institutions' model

We also estimated an 'institutions only' form of (11) which is motivated by Nickell et al.'s argument that LMI are basically enough to explain the evolution of unemployment in the OECD (Nickell et al., 2005, p. 22). Figure 12 in the appendix depicts the results for the estimated coefficients of the tax wedge variable for this parsimonious model that could be estimated with the full sample. The results do not deviate substantially from the 'shocks-and-institutions' model, but the estimated tax wedge coefficients are even more negative. The heterogeneity of the 'LMI-only' model is at least as large as before. There is again evidence for a significant divergence in the coefficient of interest over the range of countries.

Against the results presented thus far it is hard to perceive that a panel model with common slopes is supported by a test for poolability. This is confirmed by applying the Roy-Zellner test to (11) as described for instance by Baltagi (2008). The null of $H_0 : \beta_i = \beta$ for all i is clearly rejected ($\chi^2(192) = 1366.42$).[12]

4.3.2 Assessing congruency

Before we consider the possibility of simplifying the SUR-model we need to test for misspecification, or in the *Gets* terminology assess our GUM's congruency (Campos et al., 2005). The most worrisome aspect of misspecification is certainly nonsphericality of the errors, and auto-

[12] consider Swamy's test here

correlated disturbances in particular.[13]

A simple way to test for $H_0 : cov(\varepsilon_{i,t}, \varepsilon_{i,t-1}) = 0 \quad \forall \quad i$, based on the LM-princple, is to run an artificial regression of the estimated SUR residuals on their lagged values and all of the regressors included in (11). The standard t-test on the coefficient of the lagged residuals is then used.[14]. If we do this for each country, running the SUR model in first differences, the test does reject H_0 in three cases, Austria, Germany, and Japan.[15]

Even if the presence of autocorrelated disturbances at this early point of the model selection process seems almost negligible, we should not, from the *Gets* approach, content ourselves too easily with the simple dynamic structure of (11). One possible explanation for the mainly negative coefficients of the tax wedge could be a poor approximation to real world dynamics. In particular one might hypothesise that there is a certain time lag between the change in taxation and its outcome on the labour market. To explore this possibility we add two lags of the tax wedge variables and analyse the joint impact of changes in taxation. However, before we carry out the estimations we want to reduce the number of right-hand-side variables in order to gain additional degrees of freedom.

Before we proceed, we want to refer quickly to the practice in panel models in the respective literature. The contributions by Nickell et al. (2003) and Nickell et al. (2005) confirm that in fixed-effects panel models − with the data in levels − the estimated errors are not white noise. This can be easily concluded from the fact that both papers make extensively use of (feasible) generalised least squares (GLS). In our dataset this is easily confirmed by applying Wooldridge's test (Wooldridge, 2002) for autocorrelation to the data in levels.[16] Other tests and even simple visual inspections of the residuals support this conclusion.

4.3.3 Model reduction

An explorative analysis of the correlation structure among the LMI and the macroeconomic variables suggests that there could be some redundancy in the set of regressors. More particular employment protection and bargaining coordination show a rather high degree of positive correlation for low values of both variables. For the macro shock variables positive correlations among output gap, TFP and labour demand shocks are detected.

Reducing the set of regressors systematically, one at a time, does not lead to smaller values of the model selection criteria used − the Akaike (AIC) and the Schwarz (SIC) information criterion − except for two variables. Firstly, dropping the employment protection indicators decreases both information criteria below the respective values for the GUM. The same holds for leaving out the bargaining coordination indicators. Secondly, combining both steps, i.e. dropping ep and bc dummies, reduces the AIC and SIC further.

[13] Davidson and MacKinnon (1993, p. 371) note that in dealing with multivariate regression models for time-series data, "one might well expect them to display serial correlation". Beck (2001, p.3) points out that "[i]t is unlikely that cross-national panel errors will meet the assumption of sphericality".

[14] This kind of artificial regression is also called Gauss-Newton-Regression, see Davidson and MacKinnon (1993, pp. 176-208) for further details

[15] The problem of autocorrelated disturbances is more severe for the model in levels.

[16] We make use of Stata's xtserial routine.

The reduced model that results from this selection strategy does not have the nice properties of the full model when it comes to assessing the residuals. For six countries we find significant signs of autocorrelated residuals. This result points, similiarly as the tests for the full model, to potentially ill-specified dynamics of the current model.

4.3.4 Lag truncation

The lag structure – or lag truncation – is clearly part of the model selection process. As already mentioned before the impact of taxation, as well as other variables, could extend beyond the contemporaneous period. Adding the second lag of the dependent variable to (11) and including first and second lags of the tax wedge in the matrix of regressors and continuing with the reduced model, i.e. without ep and bc, leads only to a marginal change in the results of the LM-test for serial correlation. The tests show that the residuals are still plagued by autocorrelation (level of significance $\alpha = .05$) in the equations for Australia, Austria, Japan, the UK and the US.[17] Surprisingly the US is new to the list, hinting at the possibility that the added lags might introduce serial correlation themselves. Conventional F-tests for the joint significance of the lagged dependent variable fail to reject the null for Austria, the UK, and the US amongst others. The tax wedge and its lags turn out significant for all but two countries, Japan and the Netherlands.

We take these findings as an indication to give up the 'one-size-fits-all' approach at this point. Thus in the following the uniform choice of the set of regressors for all countries is no longer maintained. In particular we drop the lagged dependent variable for the US, and the second lag of it for Australia, Austria, Japan and the UK. For these four countries we also add lags for the TFP-shock variable. These measures are sufficient to finally get rid of serially correlated residuals in all but one case. For the UK further effort is needed. Eventually there is also a specification for the UK, including also the terms-of-trade shock and the real interest rate in first lags, that leads to nonspherical errors.

The coefficients for the final SUR model with country-specific equations are displayed in Table 13 in the appendix. Here, we report the distribution of the estimates for the long-run multiplier (Figure 2):

$$\phi = \frac{\hat{\beta}_{\Delta tw} + \hat{\beta}_{\Delta tw_{-1}} + \hat{\beta}_{\Delta tw_{-2}}}{1 - \hat{\beta}_{\Delta ur_{-1}} - \hat{\beta}_{\Delta ur_{-2}}}$$

These numbers are rather symmetrically distributed with a mean of 3. Only three of the long-run multipliers are significantly different from zero at a level of significance .05: Austria (-22.7), Norway (-7.9), and the US (24.5). The estimates are still quite heterogeneous as the 'fat-tailed' distribution demonstrates. With around 40% of the estimates below zero the overall picture about the influence of taxes on unemployment remains rather unclear.

[17] The Breusch-Pagan test of independence leads again to the rejection of the null, i.e. there is cross-section correlation.

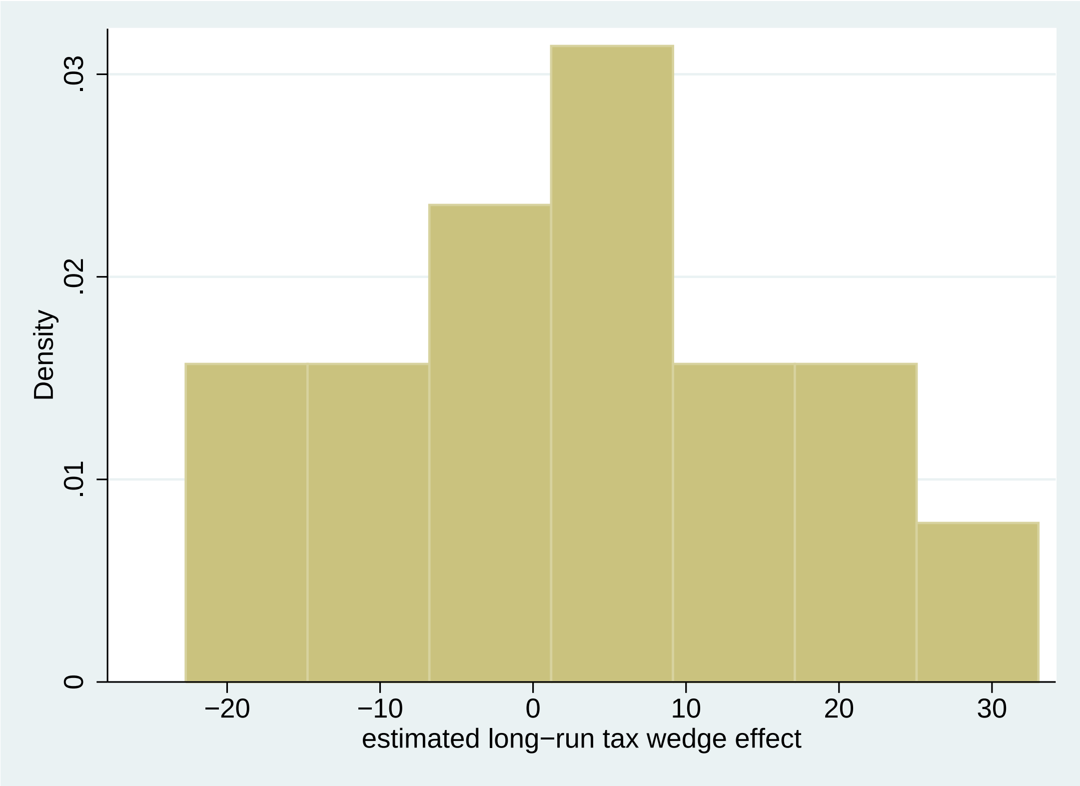

Figure 2: Distribution of long-run multiplier from SUR model

4.3.5 Nonlinearity

In section 2 signs of nonlinearity appeared when we plotted unemployment against the tax wedge. A more formal way to explore this issue is so to look at the patterns of correlation for the pooled sample and by country. Simply squaring the tax wedge into a new variable tw^2 is not very helpful since the correlation between tw and tw^2 is about .99 resulting in almost identical coefficients of correlation with the unemployment rate. If we use the standardised tax wedge instead (z_{tw} with $\mu = 0$, $\sigma = 1$), we find a negative and significant coefficient of correlation (between ur and tw^2) of $-.095$ for the pooled observations. The country-specific coefficients range between $-.81$ for Spain and .95 for France.

Admittedly, these figures are of very limited use unless tw and tw^2 are analysed together. Since the coefficient of determination of a multiple regression can be interpreted as the multiple correlation coefficient we run simple OLS models and study the $\bar{R}^2$ of these. The estimates for the pooled sample with country fixed-effects point to a hump-shaped pattern with a positive coefficient of tw and a negative for tw^2. The $\bar{R}^2$ equals .34. There are rather different paradigms behind these averages that we discover by running a SUR model, regressing ur on tw and tw^2. One group of countries displays two positive coefficients. That would mean a more than proportional effect of taxes on unemployment. There are also examples of two negative coefficients.

The results differ sharply if we look at changes in the tax wedge (Δtw and Δtw^2). For the pooled sample we obtain two negative coefficients. The joint correlation with Δur as measured by the adjusted R^2 is -.02. This bleak result is confirmed by the country-specific estimates. We can thus conclude that once we move to the differenced data, there is not much in the (cor)relation between unemployment and the tax wedge left. The two countries with the highest $\bar{R}^2$, Australia (.17) and Norway (.11), display negative coefficients for the standardised first differences of Δtw and Δtw^2. Based on these findings we do not pursue the possibility of nonlinearity any further. Such an endeavour is pointless unless one can use the data in levels.

4.4 Homogeneous panels – Pooling the data anyway?

It has been argued that pooling heterogeneous data should be preferred to the estimation of single (country) models under certain circumstances (Baltagi et al., 2000). The reason behind this argument is that no reliable estimates can be found for models fitted to individual series, if the variability of the individual time series is rather large. Baltagi et al. (2000) demonstrate this line of thinking for data of the demand for cigarettes in 46 states of the US. There is also support from Baltagi and Griffin (1997) based on an international gasoline-consumption panel.

The advantage of common-slopes panel models, however, depends either on the (theoretical) restrictions a modeller is willing to accept – as in the empirical examples just mentioned – or on the relative performance of such panel models in forecasting (see e.g. Baltagi and Griffin, 1997). In our case this would mean that there should be strong reasons to expect taxes to impact on unemployment in all countries of the sample in the same way. We doubt that our application is comparable to the demand for cigarettes or gasoline. Here, a positive effect of taxes as well as neutrality could be expected.

Nevertheless, we estimate the kind of panel models which are common in the relevant literature. The most widespread type of panel model assumes country-specific fixed effects and common slopes for all covariates (see e.g. Nickell et al., 2005). Under the *Gets* strategy this is obviously a large jump to a specific modeling class. Hence we are also going to consider solutions in between, i.e. heterogeneous panels and mixed models, in the proceeding section.

Ignoring misspecification issues for the moment, the fixed-effects model including the lagged dependent variables and the tax wedge up to its second lag yields an estimate for the contemporaneous coefficient of -1.59 (std. err. 1.23).[18] The long-run multiplier is 2.17 with a large standard error of 3.11. The residuals of this regression most likely contain an AR(1) process as indicated by the Wooldridge test and the LM-type test based on a Gauss-Newton regression.

If we estimate the model by feasible generalised least squares (FGLS), using Stata's *xtgls* command, we obtain an estimate for the contemporaneous tax wedge of 0.09 (std. err. .84) and a long-run multiplier of 3.24 (std. err. 1.78).[19] These estimates are based on a restricted sample of 16 countries and a shortened observation period of 1969 to 2009.

Using Stata's *xtregar* command, which can deal with unbalanced panels, we find an estimate of -1.34 (std. err. 1.18) and a very poorly estimated long-run multiplier. So there is a sign

[18] For the complete set of estimates see Table 13 the appendix.
[19] For the complete set of estimates see Table 13, column 2 in the appendix.

	ϕ	std. err.	p-val.
Northern Europe	-4.07	5.26	0.44
Southern Europe	-.14	8.52	.99
Anglo-Saxon	14.01	6.27	.03

Table 5: Group-specific estimates

change in our parameter of interest if we replace one FGLS estimator by another. That difference is neither due to the shorter time series in the *xtgls* procedure, as we checked by restricting the observations in the *xtregar* estimates accordingly, nor can it be explained by country-specific coefficients (ρ) in the AR(1)-errors. Assuming a common AR(1) coefficient results in a small estimate (-.04) while the interval of the country-specific estimates is defined by [-0.30;0.33].

Combining the results of the common-slopes test, the SUR estimates and the FGLS models for homogeneous panels with serially correlated errors, we conclude that there is little to no evidence that justifies homogeneous panel models. There are no hints that the common-slopes approach with stationary data could be plagued by data problems such as multicollinearity. It simply seems as if bunching 16 or 20 OECD countries together into one model is not a good modelling strategy.

4.5 Sample split

A rather simple response to the detected heterogeneity and the doubts remaining with homogeneous panels is the division of the N countries of the complete sample into subgroups. This could either be based on multivariate methods like cluster analysis or on more informal judgements, taking on an overall assessment of the institutional set-up of a candidate country. Following the approach of Eichhorst et al. (2010) we distinguish three clusters of OECD economies:

- 'Northern Europe': Austria, Belgium, Denmark, Finland, Germany, the Netherlands, Norway, Sweden, Switzerland
- 'Southern Europe': France, Italy, Portugal, Spain
- 'Anglo-Saxon': Australia, Canada, Ireland, Japan, New Zealand, United Kingdom, United States

The results of fixed-effects panel models with AR(1) error terms are summarised in Table 5. This exercise reveals different magnitudes of the coefficient of interest. The amount of heterogeneity found for the split sample is – as expected – less than in the SUR model. There is only a significant long-term effect for the 'Anglo-Saxon' group, which is remarkable in two aspects. Theoretically we would have expected to find a significant influence of labour taxes in countries where collective bargaining is the dominant way in which wages are set. From the SUR estimates we may conclude that the only country which fits into this pattern is Belgium.

With Austria, Germany, and the Nordic countries exhibiting no or even negative effects of labour taxes on unemployment in the SUR model, it is not very surprising that grouping these economies together does not lead to a significant and positive estimate.

4.6 Heterogeneous panel models

A different approach to heterogeneity is provided by the mean-group (MG) estimator (Pesaran and Smith, 1995) and the pooled mean-group (PMG) estimator (Pesaran et al., 1999). The MG estimator is simply the average of the N individual estimates, either weighted or unweighted. The PMG is an intermediate estimator with two types of coefficients: short- and long-run. The method is best explained by denoting the basic model in its error-correction form

$$\Delta y_{it} = \alpha_i + \beta_i \Delta x_{it} + \lambda_i \left(\theta x_{i,t-1} - y_{i,t-1}\right) + u_{it} \tag{12}$$

with $u_{it} \sim N\left(0, \sigma_i^2\right)$.

(12) is a special case of an autoregressive distributed lag (ADL) model (Davidson and MacKinnon, 1993, p. 682f). ADL models can be rewritten in many different ways. The error-correction form is actually only one possibility amongst several. Economically the specification of the error-correction term $\lambda_i(\theta x_{i,t-1} - y_{i,t-1})$ is crucial, because this part of the model postulates an equilibrium relationship between the dependent and the independent variables. The short-run dynamics are typically determined by statistical aspects.

The difference between the MG and the PMG estimator is that the latter assumes a common relationship (θ) between the dependent (y) and the independent variables (x) for all units i in the error-correction term, i.e. $\theta_i = \theta \; \forall i$. In the short run unit specific effects (α_i) and parameters (β_i) are allowed. Also the speed of adjustment (λ_i) can differ across units.

The PMG estimator can handle both, stationary as well as a nonstationary time series (Pesaran et al., 1999). For this reason we also run a model in levels, third column in Table 6. Note that in the table the models in first differences (columns 1 and 2) and the model in levels are presented together, hence the parentheses around the difference operator (Δ).

The MG estimatates of the tax wedge are all insignificant. This holds whether only the contemporaneous tax wedge or additional lags are contained in the model. For all the other LMI variables no significant effect can be found.

The PMG results do not differ much in qualitative respect. In the basic empirical model corresponding to (12) we assume that only tax wedge, benefit replacement rate and trade union density enter the long-run relationship. The macroeconomic (shock) variables are all assumed to be transitory and thus country-specific. The tax wedge turns out to be insignificant. Only trade union density has a significant and correctly signed effect on unemployment.

The MG and PMG estimators allow for the possibility to test the restriction in the PMG model of common slopes in the error-correction term see (see Blackburne III and Frank, 2007). This is a Hausman-type specification test with a consistent estimator (MG) under the null and the alternative hypothesis. The PMG estimator is only consistent under H_0, and also efficient.

	(MG) $\Delta^2 ur$	(PMG) $\Delta^2 ur$	(PMG) Δur
		Long-run coefficients	
$(\Delta)tw$	1.663	-0.918	28.92***
	(0.27)	(-0.28)	(7.60)
$(\Delta)brr$	-0.0725*	-0.0493**	-0.177***
	(-2.15)	(-2.73)	(-5.33)
$(\Delta)tud$	0.109	0.192***	0.126***
	(1.78)	(5.44)	(3.35)
		Short-run coefficients	
$\bar{\lambda}$	-0.512***	-0.471***	-0.0652***
	(-13.98)	(-15.55)	(-3.33)
$\Delta^{(2)}ur_{-2}$	-0.0322	-0.0386	0.255***
	(-0.99)	(-1.27)	(4.60)
$\Delta^{(2)}tw$	-0.642	-0.911	-4.386*
	(-0.30)	(-0.71)	(-2.29)
$\Delta^{(2)}tw_{-1}$	0.121	-0.218	-3.760
	(0.10)	(-0.23)	(-1.66)
$\Delta^{(2)}brr$	0.0150	0.0176	0.00918
	(0.95)	(1.03)	(0.33)
$\Delta^{(2)}tud$	0.0350	0.0306	0.0796*
	(0.58)	(0.57)	(2.37)
tfpsh	-29.95***	-29.40***	-33.15***
	(-6.59)	(-6.85)	(-5.17)
out	-16.28***	-15.42***	-20.38***
	(-8.13)	(-8.33)	(-7.52)
lds	9.379*	8.568*	23.36***
	(2.32)	(2.34)	(4.10)
tts	0.845	0.752	1.528
	(0.80)	(0.70)	(0.86)
rirl	2.103	2.231	4.963*
	(1.02)	(1.36)	(2.56)
N	874	874	875

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Pooled Mean Group model

The test statistic, which is $\chi^2(3)$ distributed, is calculated as 2.71. The corresponding p-value is .44 leading to the conclusion that the restriction of common slopes in the error-correction term is not rejected by the data and the PMG estimator should thus be preferred.

Eventually, the PMG estimator can also be used with the data in levels. Assuming that only unemployment, taxes, replacement rates, and union density are cointegrated we estimate a model in levels with the macroeconomic variables included in the short-run dynamics. Surprisingly a significant positive long-run effect of the tax wedge on unemployment shows up in this approach. There is, however, also a significant negative short-run effect. Jointly Δtw and Δtw_{-1} are different from zero at a p-value of .07.

Abstracting from the fact that the long-run coefficient β^{lr}_{tw} is actually the same for all countries while the short-run coefficients β^{sr}_{tw} and $\beta^{sr}_{tw(-1)}$ in Table 6 are averages of country-specific estimates, we can note that

$$\beta^{lr}_{tw} = -\frac{\tilde{\beta}^0_{tw}+\tilde{\beta}^1_{tw)}}{\lambda}$$

The two coefficients on the right-hand-side of this expression are the same as the 'level' coefficients in the following representation of the ADL model:

$$y_{it} = \alpha_i + \delta y_{i,t-1} + \beta^0 x_{it} + \beta^1 x_{i,t-1} + u_{it} \tag{13}$$

If one is willing to approximate $\tilde{\beta}^1_{tw}$ by the estimate for Δtw in column 3 of Table 6, then $\tilde{\beta}^0_{tw}$ equals roughly 6.3.[20] Taken at face value the PMG estimator in levels provides the first sign of a significant effect of taxes on unemployment. There are two further results that cause concern. One is the wrongly signed but (highly!) significant effect of the benefit replacement rate in the long-run part of the model. Since we have not detected anything like this before, we are rather suspicious about this strange result. In the same vein the estimate for the second lag of the dependent variable (Δur_{-2}) could indicate even more serious problems. Based on the SUR estimates and confirmed by estimating the model in the form of (13) by *xtmg* in Stata (Eberhardt, 2012), it is hard to conceive that the positive coefficient is a reasonable description of unemployment dynamics.

4.7 Mixed fixed and random models

Finally we return to the so-called mixed fixed and random (MFR) models on which we touched upon when testing for (Granger) causality. This class of models is a generalisation of fixed-/random effects models in combination with random coefficients. So in general every single coefficient $(\beta_{k,i})$ can be thought of as being randomly distributed around a common mean. The distribution is assumed to be normal with mean $\bar{\beta}_k$ and variance $\sigma^2_{\beta_k}$

$$\beta_{k,i} = \bar{\beta}_k + \eta_i$$

[20] Due to the way Blackburne III and Frank (2007) set up the model in Stata, the short-run coefficient estimate corresponds to the lagged variable in (13).

where $\eta_i \sim N(0, \sigma_\eta^2)$.

The econometrician faces the problem to determine the stochastic structure of the model. Two alternative model selection strategies appear appropriate. On the one we could start with a large set of random coefficients and eliminate, possibly step by step, those variables where there is little variation across countries. On the other hand there might be a priori reasons to restrict the set of randomly distributed coefficients. We opt for the second way and assume that the 20 economies of our sample simply differ in the speed of adjustment. Technically speaking we model the coefficients of the lagged dependent variable (first and second lag) as random. In addition we also run a model in which only the coefficient of the tax wedge is treated as random.

In the process of model selection we encountered unprecedented difficulties for the macro variables. For this reason we decided to replace TFP-shock, output gap and labour demand shock by the scores from a principal component analysis. The estimates of the following model are presented in Table 7.

$$\begin{aligned}\Delta ur_{it} \quad = \quad & \alpha_i + \gamma_{1i}\Delta ur_{it-1} + \gamma_{2i}\Delta ur_{it-2} + \\ & \beta_{tw,i}\Delta tw_{it} + \beta_{brr}\Delta brr_{it} + \beta_{tud}\Delta tud_{it} + \\ & \delta_1 MACRO_{it} + \delta_2 MACRO_{it-1} + \delta_3 tts_{it} + \delta_4 rirl_{it} + \varepsilon_{it}\end{aligned}$$

In the first specification (1) $\beta_{tw,i} = \bar{\beta}_{tw} + \eta_i^{tw}$. In (2) $\beta_{tw,i}$ is assumed to be fixed, thus $\beta_{tw,i} = \beta_{tw}$. Instead γ_{1i} and γ_{2i} are considered randomly distributed around common $\bar{\gamma}_1$ and $\bar{\gamma}_2$. Finally, in (3) we combine (1) and (2).

The estimated coefficients differ only marginally in the fixed part of the model (upper part of Table 7). The tax wedge is not significant in any of the four models. Modelling its coefficient as random leads to a large, but imprecise estimate for σ_η^{tw}. In fact the standard deviation is about the same seize as the mean ($\bar{\beta}_{tw}$). This finding is in line with the results of the SUR model. This holds for the negative, but insignificant average effect, too. The magnitude of heterogeneity that we encounter in the γ_{1i} and γ_{2i} coefficients (column (3)) is rather small. This result is also largely in line with the SUR estimates. Treating the tax wedge coefficient as random does not make a difference (column (4)).

Testing the residuals of the basic MFR model (3) reveals clear signs of serial correlation in 9 of 20 countries. Adding even the third lag of the dependent variable as well as seconds lags of macro values does not seem to lessen this problem. Instead, the tests appear to indicate that for some countries our specifications are incomplete in the sense that important factors driving unemployment dynamics are missing.

5 Conclusions

'Seek and you shall find' is not the bottom line here. Although we are tempted to claim that one could actually find something nice and presentable if one seeks for long. The figures and models presented here suggest that the available macroeconomic data does not yield the empirical

	(1) Δur	(2) Δur	(3) Δur	(4) Δur
Δur_{-1}	0.524***	0.523***	0.520***	0.517***
	(17.15)	(17.12)	(11.65)	(11.42)
Δur_{-2}	−0.103***	−0.102***	−0.131***	−0.130***
	(−4.00)	(−3.99)	(−3.80)	(−3.77)
Δtw	−1.532	−1.735	−1.510	−2.105
	(−1.36)	(−1.39)	(−1.36)	(−1.32)
Δbrr	−0.00788	−0.00771	−0.00882	−0.00835
	(−0.66)	(−0.64)	(−0.74)	(−0.70)
Δtud	0.0696***	0.0695***	0.0737***	0.0732***
	(3.58)	(3.58)	(3.82)	(3.78)
$MACRO$	−0.393***	−0.393***	−0.390***	−0.391***
	(−22.76)	(−22.78)	(−22.90)	(−23.00)
$MACRO_{-1}$	0.262***	0.263***	0.261***	0.261***
	(13.21)	(13.21)	(13.18)	(13.19)
tts	2.132	2.108	2.059	2.038
	(1.69)	(1.67)	(1.67)	(1.65)
rirl	−2.630**	−2.642**	−2.510**	−2.474**
	(−3.24)	(−3.25)	(−3.15)	(−3.08)
cons	0.182***	0.183***	0.180***	0.180***
	(5.32)	(5.32)	(5.36)	(5.31)
		Random Parameters		
$\sigma_\eta(\Delta tw)$		1.344		3.619
sd		4.523		2.352
$\sigma_\eta(\Delta ur_{-1})$			0.127	0.131
sd			0.034	0.034
$\sigma_\eta(\Delta ur_{-2})$			0.088	0.089
sd			0.030	0.030
N	880	880	880	880

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Mixed Fixed and Random models

evidence we have been looking for. Again it has to be stressed that this does not mean that labour taxes are neutral to unemployment. The plain message emerging from our study is that the evidence simply is not where we looked for it. Consequently one should look elsewhere.

The heterogeneity that we encountered from the beginning should be taken as a strong sign that the effect of labour taxes on unemployment very likely differs substantially between the OECD economies of our sample. This is in line with the findings of Daveri and Tabellini (2000), Everaert and Heylen (2002), and García and Sala (2008). Unfortunately our estimates provide little new guidance for further research in this respect. There are, however, a few other ways that could be explored. One is to look at employment and wages, too. It could well be the case that the outcome variable used here, unemployment, is driven by factors which we did not cover. Another promising way could be to abandon the (implicit) assumption made here that all labour taxes impact on unemployment in exactly the same way. Technically this means treating the component of the tax wedge as separate regressors. The advances in time series analysis and in macro panel econometrics certainly provide additional technical armoury that could be used. A particular branch are factor or unobserved component models, which have been used e.g. by Berger and Everaert (2010).

The route taken by Everaert and Heylen (2002) might also be worth reconsidering. This study follows a thorough cointegration approach and comes up with significant estimates for the labour tax variable. The crucial difference to our methodological understanding is the treatment of stationary and nonstationary data. While we abstained from modelling I(1) and I(0) variables in one model, Everaert and Heylen allow this mixture. A 'third' way would be to use only macroeconomic variables which are I(1) and (panel) cointegrated with tax and unemployment rates. This would call for a different approach to data construction. Instead of using the set of variables introduced by the 'shocks-and-institutions' literature, one could follow a bottom-up approach and build up a more tailored data set.

Obviously data selection is a very important choice in empirical research. We added a piece of knowledge to the existing wisdom by exploring the use of 'micro' data in macroeconometric models. The low predictive power of these variables in our models does not necessarily mean that the data is useless in this kind of applications. Again, all we can say is that they did not turn out in a supportive sense, hence one should look either differently on them or elsewhere.

References

Arellano, M. and S. Bond (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies 58*, 277–297.

Arellano, M. and O. Bover (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics 68*, 29–51.

Baccaro, L. and D. Rei (2007). Institutional Determinants of Unemployment in OECD Countries: Does the Deregulatory View Hold Water? *International Organization 61*, 527–569.

Baltagi, B. H. (2008). *Econometric Analysis of Panel Data.* Wiley.

Baltagi, B. H. and J. M. Griffin (1997). Pooled estimators vs. their heterogeneous counterparts in the context of dynamic demand for gasoline. *Journal of Econometrics 77*, 303–327.

Baltagi, B. H., J. M. Griffin, and W. Xiong (2000). To Pool or Not to Pool: Homogeneous Versus Heterogeneous Estimators Applied to Cigarette Demand. *The Review of Economics and Statistics 82*, 117–126.

Bassanini, A. and R. Duval (2009). Unemployment, institutions, and reform complementarities: re-assessing the aggregate evidence for OECD countries. *Oxford Review of Economic Policy 25*, 40–59.

Bean, C., P. Layard, and S. Nickell (1986). The Rise in Unemployment: A Multi-Country Study. *Economica 53*, S1–S22.

Beck, N. (2001). Longitudinal (Panel and Time Series Cross-Section) Data. http://www.nyu.edu/classes/nbeck/longdata/longitude20011short.pdf (last visited: 2/29/2012).

Belot, M. and J. C. van Ours (2001). Unemployment and Labor Market Institutions: An Empirical Analysis. *Journal of the Japanese and International Economies 15*, 403–418.

Berger, T. and G. Everaert (2009). A replication note on unemployment in the OECD since the 1960s: What do we know? *Empirical Economics 36*, 479–485.

Berger, T. and G. Everaert (2010). Labour taxes and unemployment evidence from a panel unobserved component model. *Journal of Economic Dynamics and Control 34* (3), 354 – 364.

Blackburne III, E. F. and M. W. Frank (2007). Estimation of nonstationary heterogeneous panels. *Stata Journal 7*, 197–2008.

Blanchard, O. and J. Wolfers (2000). Shocks and institutions and the rise of European unemployment. The aggregate evidence. *Economic Journal 110*, C1–C33.

Blundell, R. and S. Bond (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics 87*, 115–143.

Boockmann, B., R.-H. Peters, and V. Steiner (2001). Social Security, Real Wage Resistance and Employment in Sweden, the UK and West Germany. In D. Milleker (Ed.), *Beschäftigungspolitik in Europa.* Bad Homburg: Frankfurter Institut - Stiftung Marktwirtschaft und Politik.

Calmfors, L. and J. Drifill (1988). Bargaining structure, corporatism and macroeconomic performance. *Economic Policy 3*, 13–61.

Campos, J., N. R. Ericsson, and D. F. Hendry (2005). Introduction. In J. Campos, N. R. Ericsson, and D. F. Hendry (Eds.), *General-to-specific modeling: an overview and selected bibliography*, pp. 1–81. Edward Elgar.

Daveri, F. and G. Tabellini (2000). Unemployment, growth and taxation in industrial countries. *Economic Policy 15*, 47–104.

Davidson, R. and J. G. MacKinnon (1993). *Estimation and Inference in Econometrics.* Oxford University Press.

Eberhardt, M. (2009). Nonstationary Panel Econometrics and Common Factor Models: An Introductory Reader. Unpublished mimeo, University of Oxford.

Eberhardt, M. (2012). Estimating panel time-series models with heterogeneous slopes. *Stata Journal 12*(1), 61–71.

Eichhorst, W., M. T. Feil, and P. Marx (2010). Crisis, What Crisis? Patterns of Adaption in European Labor Markets. IZA Discussion Paper, No. 5045.

Elmeskov, J., J. P. Martin, and S. Scarpatta (1998). Key lessons for labour market reforms: ecvidence from OECD countries' experiences. *Swedish Economic Policy Review 5*, 205–252.

Everaert, G. and F. Heylen (2002). Labour Taxes and Unemployment in the OECD: A Nonstationary Panel Data Analysis. SHERPPA, Ghent University, Gent, Belgium.

Feil, M. T. (2012). Taxation and unemployment - a survey. Unpublished mimeo.

García, J. R. and H. Sala (2008). The tax system incidence on unemployment: A country-specific analysis for the OECD economies. *Economic Modelling 25*, 1232–1245.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica 37*, 424–438.

Gruber, J. (1997). The Incidence of Payroll Taxation: Evidence from Chile. *Journal of Labor Economics 15*, S72–S101.

Hartwig, J. (2010). Is health capital formation good for long-term economic growth? - Panel Granger causality evidence for OECD countries. *Journal of Macroeconomics 32*, 314–325.

Hendry, D. F. (1993). *Econometrics: Alchemy or Science? Essays in Econometric Methodology.* Blackwell.

Howell, D. R., D. Baker, A. Glyn, and J. Schmitt (2006). Are Protective Labor Market Institutions Really at the Root of Unemployment? A Critical Perspective on the Statistical Evidence. CEPR Reports and Issue Briefs 2006-14, Center for Economic and Policy Research (CEPR).

Hsiao, C., D. C. Mountain, L. M. Chan, and K. Y. Tsui (1989). Modeling Ontario Regional Electricity System Demand Using a Mixed Fixed and Random Coefficients Approach. *Regional Science and Urban Economics 19*, 565–587.

Kwiatkowski, D., P. C. Phillips, P. Schmidt, and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics 54*, 159–178.

Layard, R., S. Nickell, and R. Jackman (1991). *Unemployment : Macroeconomic Performance and the Labour Market.* Oxford University Press.

Manning, A. (1993). Wage bargaining and the Phillips curve: The identification and specification of aggregate wage equations. *The Economic Journal 103*, 98–118.

Nair-Reichert, U. and D. Weinhold (2001). Causality Tests for Cross-Country Panels: a New Look at FDI and Economic Growth in Developing Countries. *Oxford Bulletin of Economics and Statistics 63*, 153–171.

Nickell, S. (1997). Unemployment and Labor Market Rigidities: Europe versus North America. *Journal of Economic Perspectives 11*, 55–74.

Nickell, S. (1998). Unemployment: Questions and Some Answers. *The Economic Journal 108*, 802–816.

Nickell, S. and R. Layard (1999). Labor Market Institutions and Economic Performance. *Handbook of Labor Economics 3*, 3029–3084.

Nickell, S. and L. Nunziata (2001). Labour Market Institutions Database, mimeo.

Nickell, S., L. Nunziata, and W. Ochel (2005). Unemployment in the OECD since the 1960s. What do we know? *The Economic Journal 115*, 1–27.

Nickell, S., L. Nunziata, W. Ochel, and G. Quinitini (2003). The Beveridge Curve, Unemployment, and Wages in the OECD from the 1960s to the 1990s. In P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford (Eds.), *Knowledge, Information and Expectations in Modern Macroeconomics*, Chapter 19, pp. 394–440. New Jersey: Princeton University Press.

OECD (1978). The tax/benefit position of selected income groups 1972-1976. Organisation for Economic Co-operation and Development, Paris.

OECD (1994). *The OECD Jobs Study: Evidence and Explanations.* OECD.

OECD (2011). *Taxing Wages 2009-2010.* OECD.

Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics 22*, 265–312.

Pesaran, M. H., Y. Shin, and R. P. Smith (1999). Pooled Mean Group Estimation of Dynamic Heterogeneous Panels. *Journal of the American Statistical Association 94*, 621–634.

Pesaran, M. H. and R. P. Smith (1995). Estimating long-run relationships from dynamic heterogeneous panels. *Journal of Econometrics 68*, 79–113.

Phillips, P. C. B. and P. Perron (1988). Testing for a unit root in time series regression. *Biometrika 75*, 335–346.

Prescott, E. C. (2004). Why Do Americans Work so Much More than Europeans? NBER Working Paper No. 10316.

Rault, C. and A.-G. Vaubourg (2011). Labour market institutions and unemployment: Does finance matter? IZA DP No. 5606.

Steiner, V. (1998). Employment and Wage Effects of Social Security Financing - An Empirical Analysis of the West German Experience and Some Policy Simulations. In J. A. Addison and P. J. Welfens (Eds.), *Labour Markets and Social Security*, pp. 319–344. Springer.

Toda, H. Y. and T. Yamamoto (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics 66*, 225–250.

Tyrväinen, T. (1995). Real Wage Resistance and Unemployment. OECD Jobs Study Working Papers No. 10.

Weinhold, D. M. (1999). A dynamic "fixed effects" model for heterogeneous panel data. unpublished manuscript, London School of Economics.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Zellner, A. (1962). An Efficient Method of Estimating Seemingly Unrelated Regressions and Tests for Aggregation Bias. *Journal of the American Statistical Association 57*, 348–368.

6 Appendix

Data

Countries in the sample
Australia (AUS), Austria (AUT), Belgium (BEL), Canada (CAN), Denmark (DEN), Finland (FIN), France (FRA), Germany (GER), Ireland (IRE), Italy (ITA), Japan (JAP), Netherlands (NEL), New Zealand (NEW), Norway (NOR), Portugal (POR), Spain (SPA), Sweden (SWE), Switzerland (SWZ), United Kingdom (UK), United States (USA)

Gross Benefit Replacement Rate (brr)
Benefit entitlement as a percentage of previous earnings. Here the average of the gross unemployment benefit replacement rates for two earnings levels, three family situations and three durations of unemployment is used.
Source: OECD's Benefits and Wages (2009).

Trade Union Density (tud)
This variable is constructed as the ratio of total reported union members (less retired and unemployed members).
Sources: Labour Market Institutions Database by Nickell and Nunziata (2001), ICTWSS database.

Co-ordination Index (bc)
This captures the degree of consensus between the actors in collective bargaining. 1 low, 5 high.
Source: ICTWSS database.

Employment Protection Index (ep)
This captures the strictness of employment protection laws. 0 low, 2 high.
Sources: Labour Market Institutions Database by Nickell and Nunziata (2001), OECD.

Total Tax Wedge (tw)
This variable is the sum of the payroll tax rate, income tax rate and consumption tax rate.

$$\begin{aligned} \tau_{PR} &= \frac{EC}{IE - EC} \\ \tau_I &= \frac{T^I + WC}{HCR} \\ \tau_C &= \frac{T^C - S}{CC} \end{aligned}$$

EC = employers' private and public social security contributions, IE = compensation of employees, T^I = income taxes, WC = workers' social security contributions, HCR = households' current receipts, T^C = indirect taxes, S = subsidies, CC = final consumption expenditure.
Sources: Nickell and Nunziata (2002). Own calculations using OECD data.

Unemployment Rate (ur)
OECD standardised unemployment rates. ILO definition.

Terms of Trade Shock (tts)
The terms of trade shock is the change in the log of real import prices times the share of imports in GDP. Real import prices are defined as the import price deflator normalised

on the GDP deflator.
Own calculations using OECD data.

Real Interest Rate (rirl)
Long term nominal interest rate less the current rate of inflation from the OECD Economic Outlook Database.

Total Factor Productivity Shock (tfpsh)
Based on the Solow residual for each country, smoothed using a HP filter. tfpsh then is the deviation of the Solow residual from its HP filter trend.

Labour Demand Shock (lds)
Residuals from country specific employment equations, each being a regression of employment on lags of employment, real wages and output. Constructed in the same way as in Nickell and Nunziata (2001).

Output Gap (out)
Deviation of the real GDP from its HP filter trend. Own calculations using OECD data.

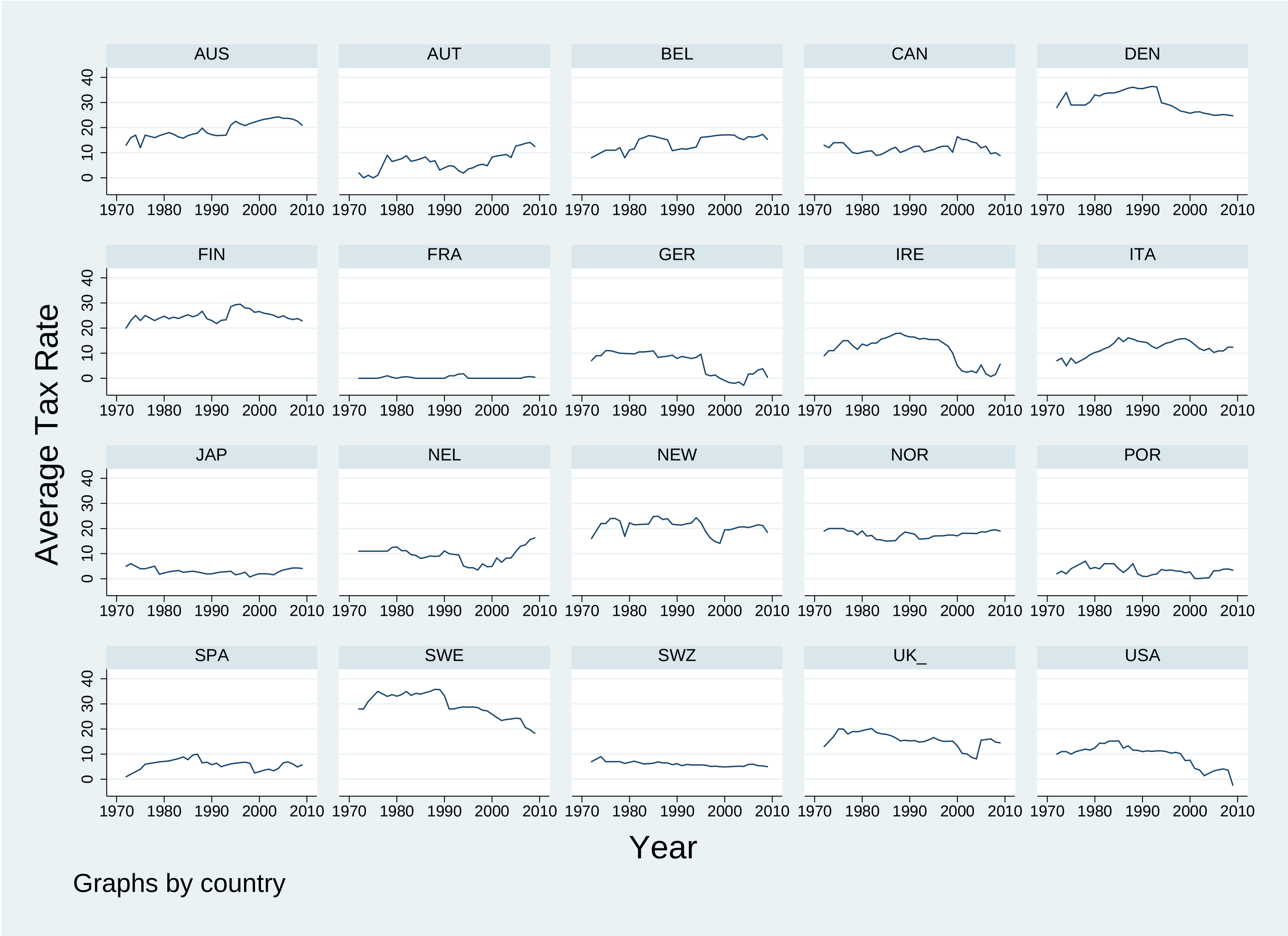

Figure 3: Average income tax rates (1972-2009)

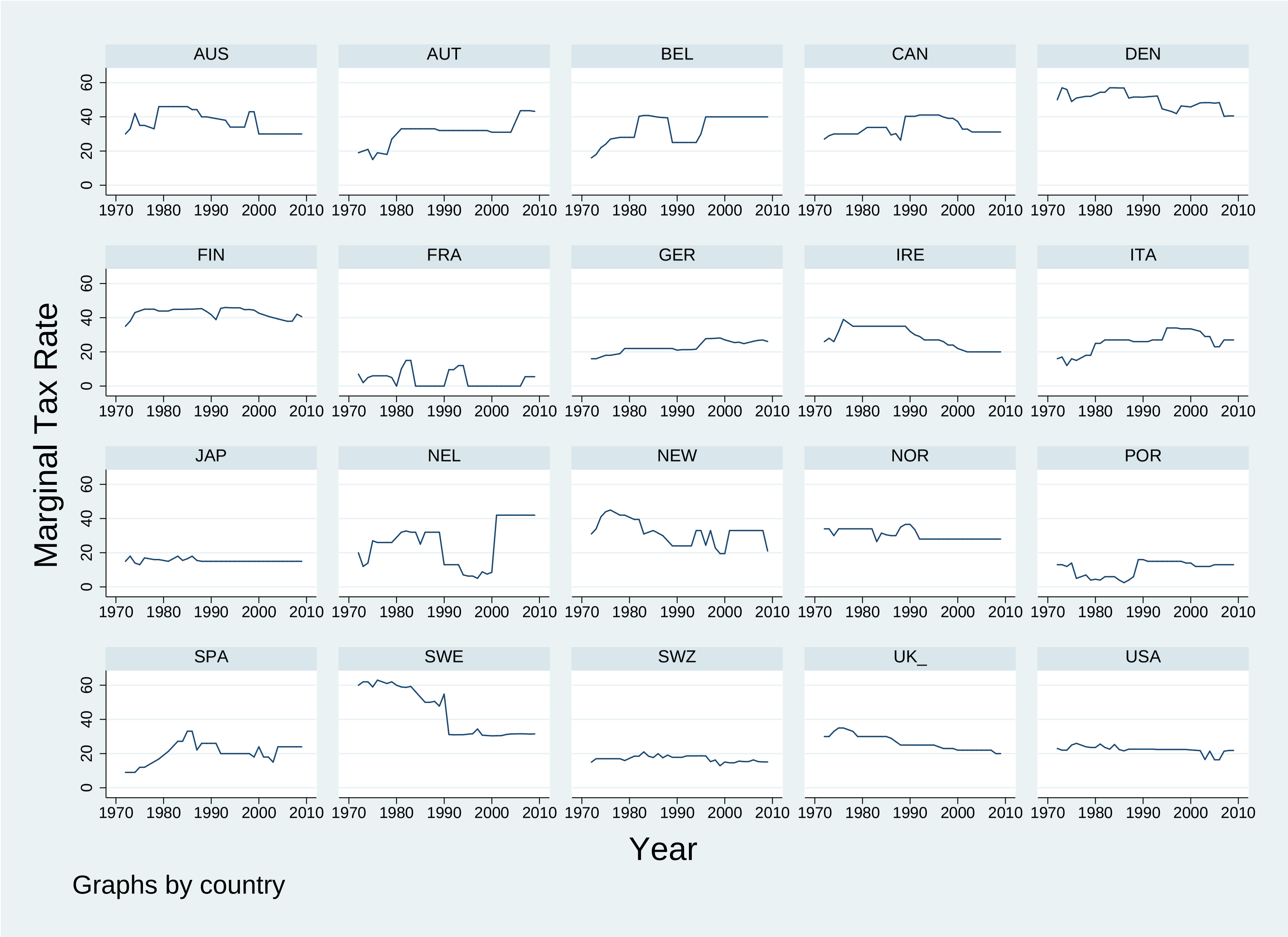

Figure 4: Marginal income tax rates (1972-2009)

Figure 5: Country-specific correlations between average tax rate and unemployment rates

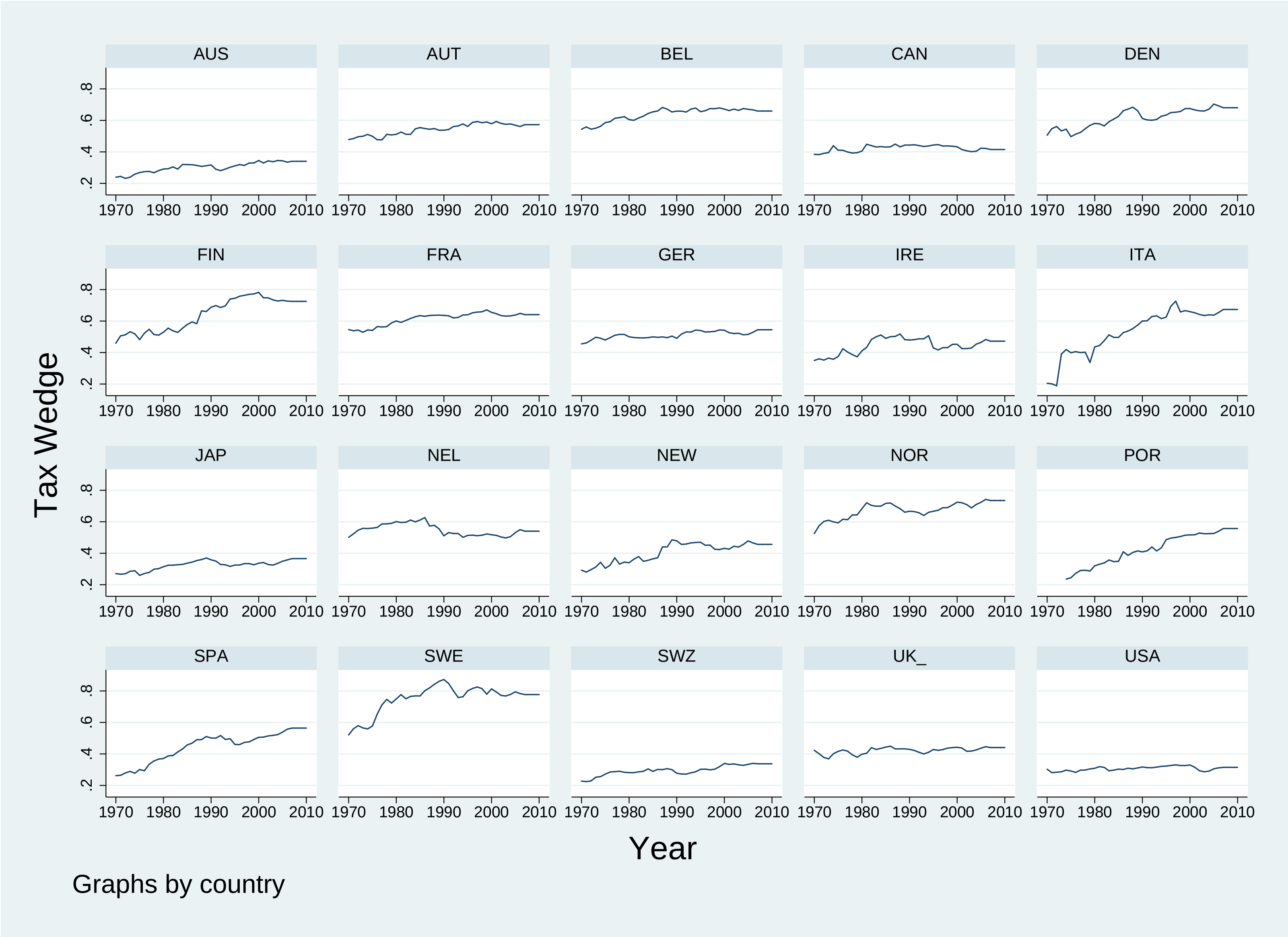

Figure 6: Tax wedge rates

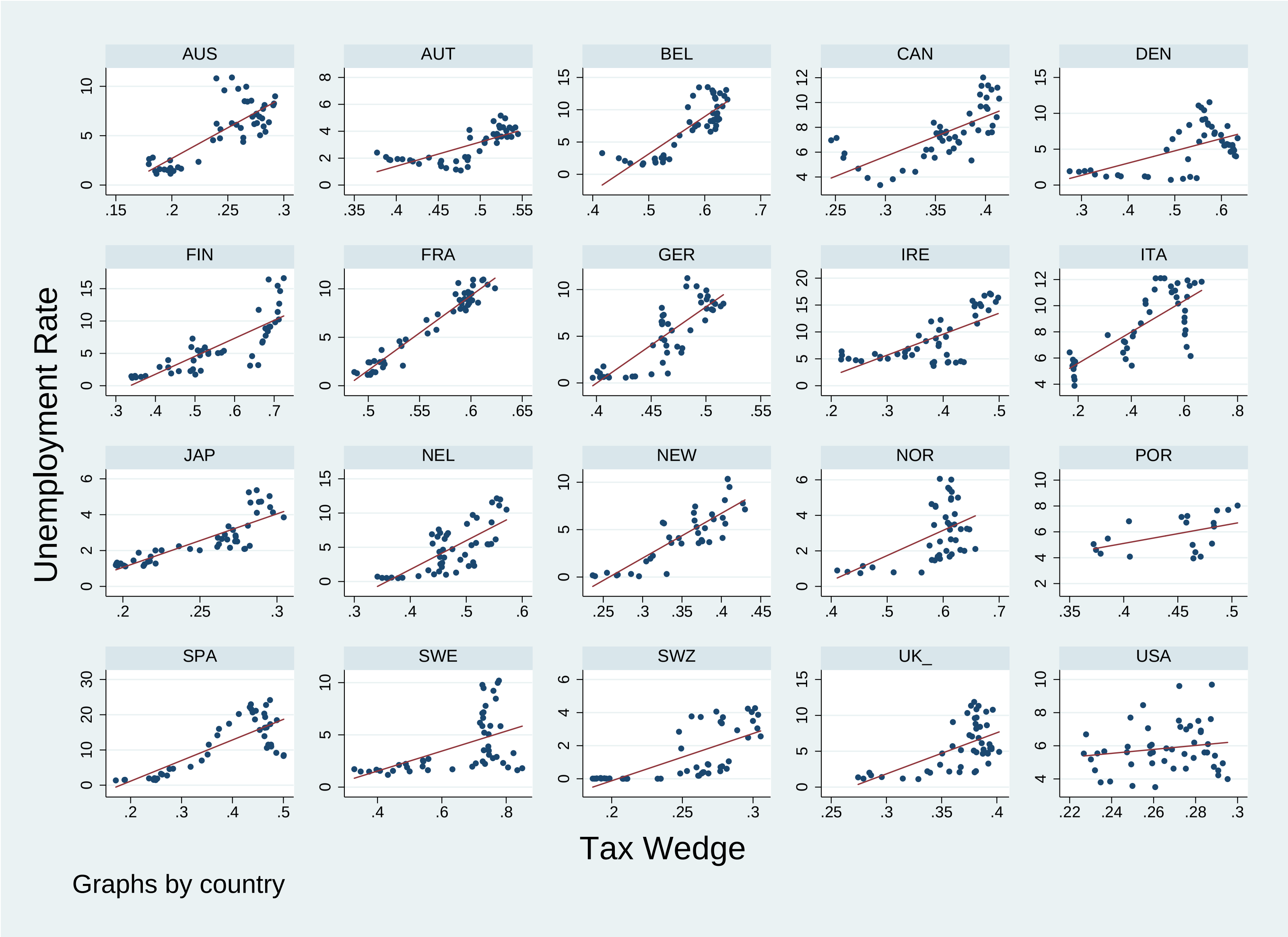

Figure 7: Country-specific correlations between tax wedge and unemployment rates

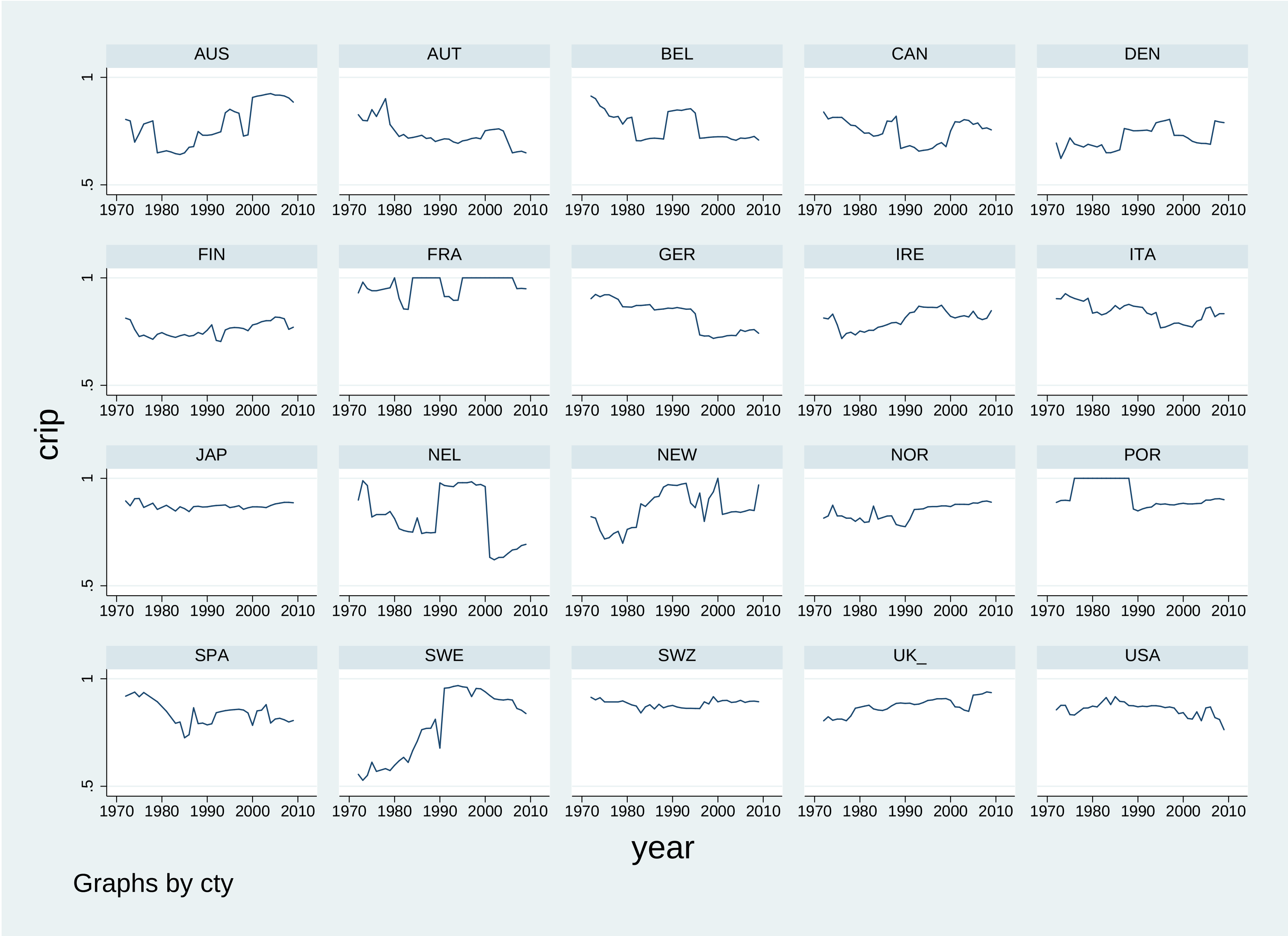

Figure 8: Coefficient of residual income progression (CRIP)

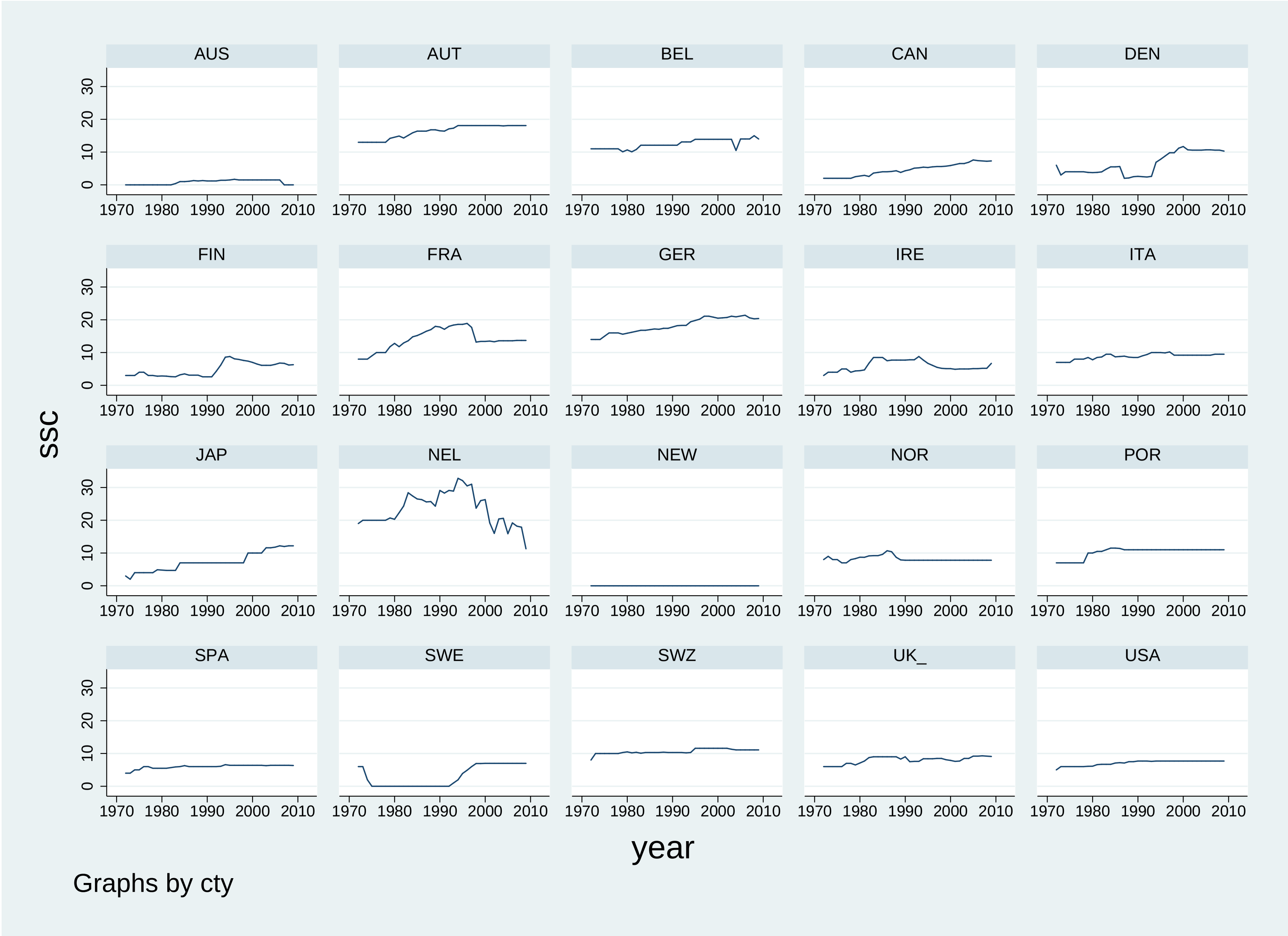

Figure 9: Social security contribution rates

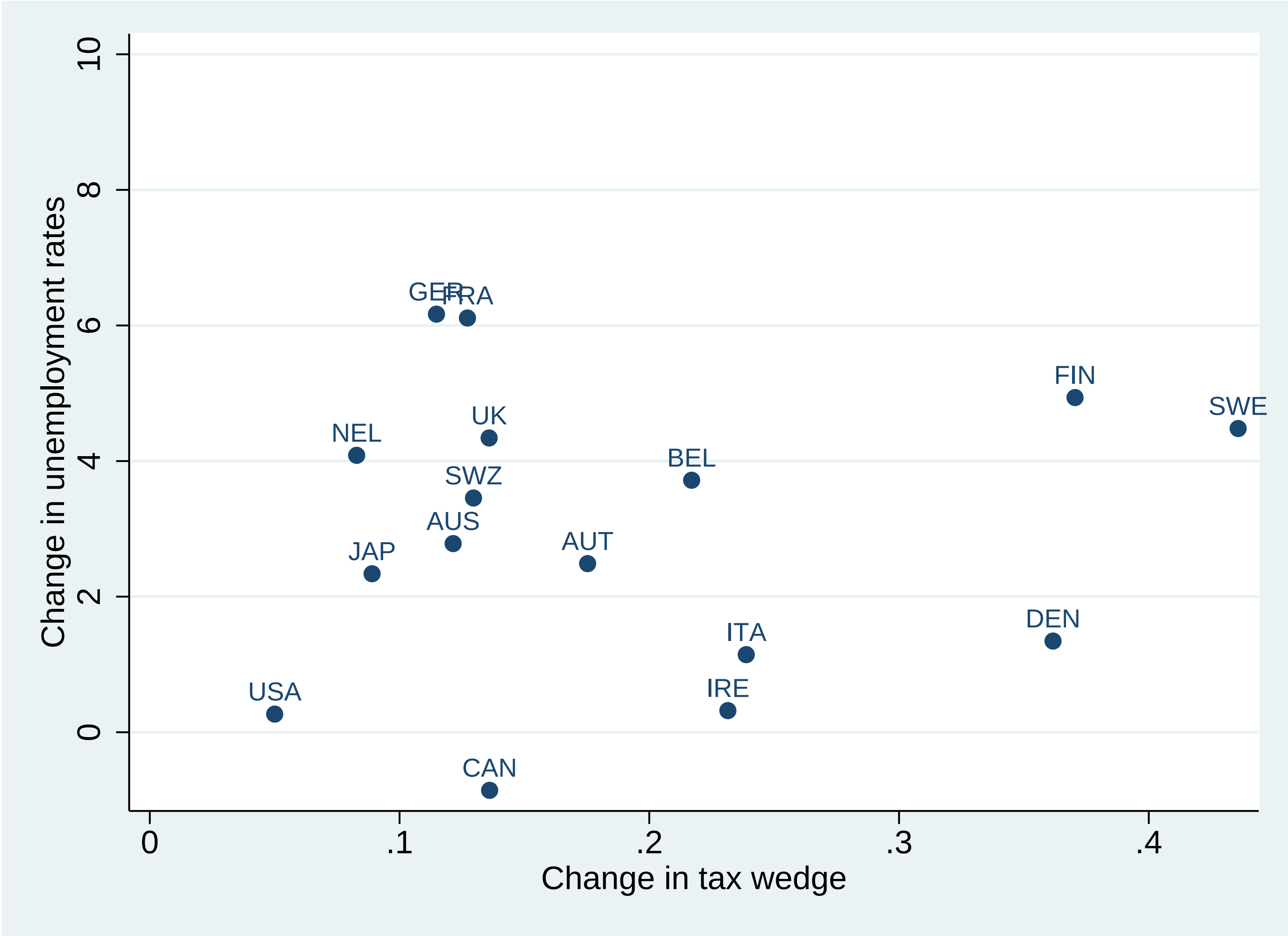

Figure 10: Change in unemployment and tax wedge (1960-2008)

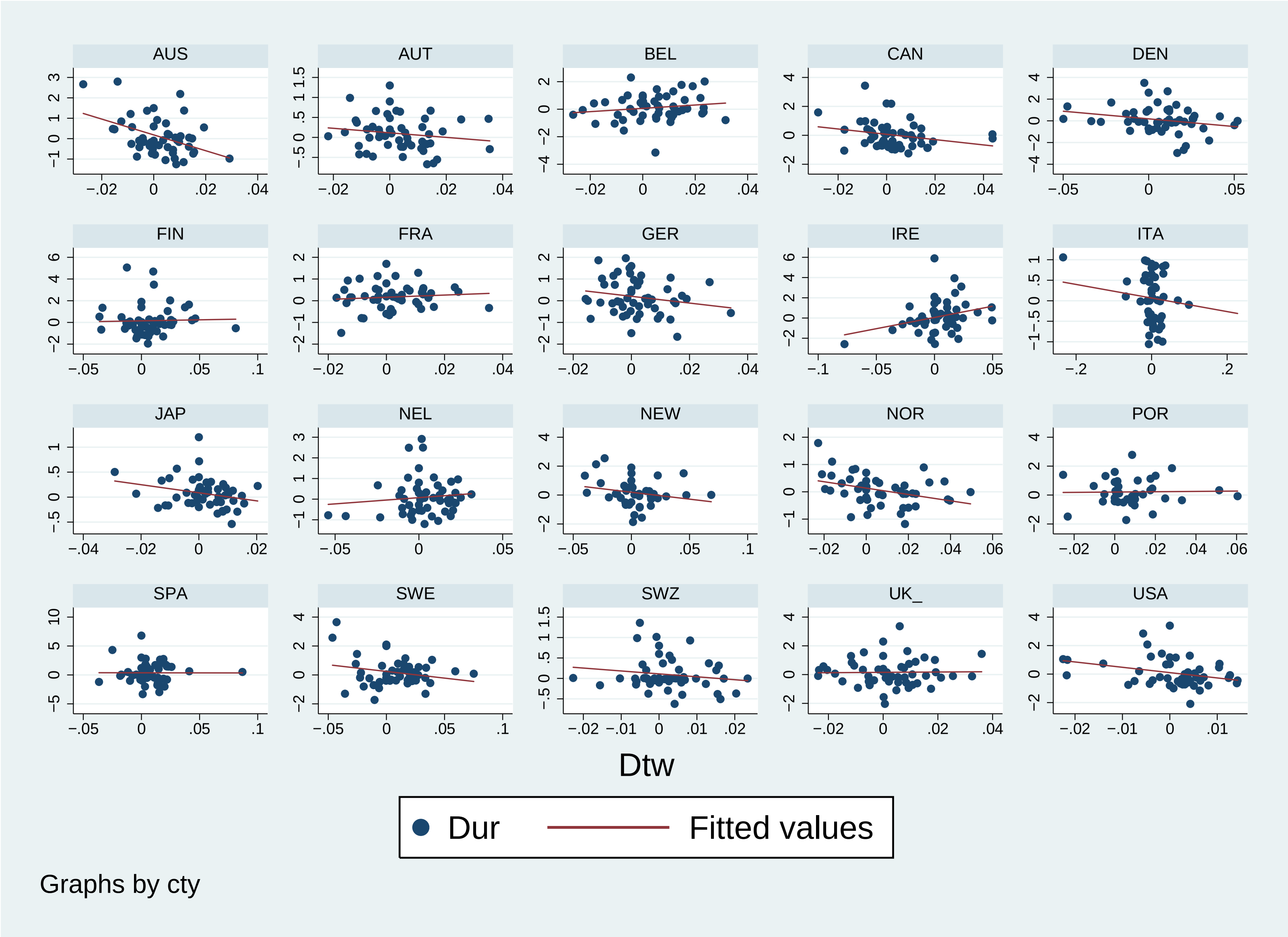

Figure 11: Changes in unemployment and tax wedge by country

country	pval trend	trend	pval τ_trend	pval drift	drift	pval τ_drift	pval τ	outcome
AUS	0.97	-	0.92	0.12	-	0.92	0.92	I(1)
AUT	0.01	x	0.46					I(1)
BEL	0.28	-	0.38	0.08	-	0.44	0.88	I(1)
CAN	0.27	-	0.28	0.03	x	0.16		I(1)
DEN	0.62	-	0.61	0.06	-	0.29	0.54	I(1)
FIN	0.19	-	0.57	0.09	-	0.47	0.75	I(1)
FRA	0.24	-	0.73	0.06	-	0.58	0.57	I(1)
GER	0.00	x	0.00					I(0)
IRE	0.57	-	0.68	0.06	-	0.39	0.69	I(1)
ITA	0.70	-	0.77	0.05	x	0.38		I(1)
JAP	0.00	x	0.06					I(1)
NEW	0.94	-	0.58	0.05	x	0.23		I(1)
NEL	0.18	-	0.55	0.10	-	0.49	0.44	I(1)
NOR	0.10	-	0.26	0.04	x	0.24		I(1)
POR	0.45	-	0.26	0.01	x	0.08		I(1)
SPA	0.21	-	0.49	0.07	-	0.37	0.65	I(1)
SWE	0.00	x	0.01					I(0)
SWZ	0.00	x	0.02					I(0)
UK_	0.30	-	0.44	0.10	-	0.26	0.99	I(1)
USA	0.29	-	0.07	0.00	x	0.02		I(0)

Table 8: Augmented Dickey-Fuller tests for a unit root in unemployment rates

country	pval trend	trend	pval τ_trend	pval drift	drift	pval τ_drift	pval τ	outcome	
AUS	0.13	-	0.66	0.13	-	0.68	0.08		I(1)
AUT	0.36	-	0.81	0.03	x	0.81			I(1)
BEL	0.69	-	0.75	0.00	x	0.04		I(0)	
CAN	0.61	-	0.62	0.00	x	0.02		I(0)	
DEN	0.19	-	0.30	0.00	x	0.03		I(0)	
FIN	0.70	-	0.98	0.03	x	0.43			I(1)
FRA	0.63	-	0.95	0.14	-	0.60	0.09		I(1)
GER	0.07	-	0.45	0.16	-	0.62	0.19		I(1)
IRE	0.61	-	0.71	0.02	x	0.19			I(1)
ITA	0.13	-	0.69	0.09	-	0.70	0.21		I(1)
JAP	0.07	-	0.55	0.26	-	0.76	0.24		I(1)
NEW	0.21	-	0.59	0.01	x	0.19			I(1)
NEL	0.39	-	0.73	0.03	x	0.29			I(1)
NOR	0.29	-	0.26	0.00	x	0.04		I(0)	
POR	0.08	-	0.53	0.00	x	0.26			I(1)
SPA	0.44	-	0.88	0.01	x	0.35			I(1)
SWE	0.97	-	0.85	0.01	x	0.12			I(1)
SWZ	0.09	-	0.54	0.13	-	0.64	0.12		I(1)
UK_	0.06	-	0.04	0.00	x	0.04		I(0)	
USA	0.04	x	0.05					I(0)	

Table 9: Augmented Dickey-Fuller tests for a unit root in tax wedge series

Country		Lag truncation parameter (l)					
		0	1	2	3	4	5
		η_μ: 5% critical value is 0.463, η_τ: 5% critical value is 0.146					
AUS	level	2.667	1.384	0.954	0.738	0.608	0.521
	trend	0.974	0.510	0.357	0.280	0.234	0.204
AUT	level	4.154	2.215	1.564	1.233	1.028	0.887
	trend	0.310	0.183	0.144	0.126	0.115	0.108
BEL	level	2.828	1.445	0.989	0.765	0.632	0.544
	trend	0.940	0.484	0.335	0.263	0.221	0.194
CAN	level	1.647	0.866	0.611	0.488	0.414	0.365
	trend	0.750	0.399	0.286	0.232	0.200	0.179
DEN	level	1.811	0.950	0.665	0.524	0.440	0.384
	trend	0.864	0.457	0.321	0.255	0.216	0.191
FIN	level	3.182	1.653	1.155	0.913	0.769	0.675
	trend	0.319	0.170	0.124	0.103	0.093	0.087
FRA	level	4.340	2.225	1.518	1.164	0.951	0.809
	trend	0.954	0.493	0.341	0.266	0.222	0.194
GER	level	4.688	2.419	1.663	1.285	1.056	0.902
	trend	0.319	0.179	0.139	0.124	0.120	0.120
IRE	level	1.004	0.523	0.364	0.286	0.240	0.210
	trend	0.768	0.398	0.276	0.217	0.182	0.159
ITA	level	2.721	1.388	0.947	0.728	0.599	0.514
	trend	0.981	0.504	0.347	0.270	0.225	0.196
JAP	level	4.486	2.318	1.597	1.238	1.021	0.876
	trend	0.430	0.236	0.175	0.146	0.132	0.124
NEL	level	1.459	0.748	0.514	0.399	0.330	0.285
	trend	1.047	0.536	0.368	0.285	0.237	0.205
NEW	level	3.471	1.789	1.231	0.954	0.788	0.677
	trend	0.679	0.355	0.249	0.197	0.167	0.148
NOR	level	3.353	1.739	1.208	0.944	0.786	0.680
	trend	0.538	0.290	0.210	0.172	0.151	0.139
POR	level	0.366	0.211	0.161	0.140	0.130	0.126
	trend	0.348	0.197	0.149	0.128	0.118	0.113
SPA	level	3.042	1.562	1.072	0.830	0.686	0.591
	trend	0.860	0.441	0.304	0.237	0.198	0.173
SWE	level	3.579	1.880	1.324	1.049	0.884	0.775
	trend	0.248	0.135	0.100	0.085	0.078	0.075
SWZ	level	4.479	2.320	1.602	1.243	1.026	0.880
	trend	0.546	0.298	0.221	0.186	0.167	0.156
UK	level	2.255	1.171	0.813	0.636	0.530	0.460
	trend	0.871	0.450	0.313	0.245	0.206	0.180
USA	level	0.428	0.242	0.185	0.161	0.147	0.138
	trend	0.324	0.182	0.140	0.121	0.111	0.104

Table 10: KPSS tests for stationarity of unemployment rates

Country		Lag truncation parameter (l)					
		0	1	2	3	4	5
		η_μ: 5% critical value is 0.463, η_τ: 5% critical value is 0.146					
AUS	level	4.732	2.446	1.671	1.284	1.052	0.898
	trend	0.566	0.326	0.238	0.194	0.167	0.150
AUT	level	4.645	2.416	1.658	1.276	1.048	0.896
	trend	0.819	0.460	0.333	0.268	0.230	0.205
BEL	level	4.316	2.271	1.570	1.217	1.005	0.863
	trend	0.990	0.548	0.393	0.315	0.271	0.241
CAN	level	2.936	1.552	1.081	0.845	0.705	0.613
	trend	1.084	0.579	0.406	0.320	0.269	0.235
DEN	level	4.264	2.211	1.526	1.185	0.982	0.848
	trend	0.875	0.465	0.331	0.266	0.229	0.206
FIN	level	4.897	2.513	1.710	1.308	1.067	0.908
	trend	0.592	0.322	0.229	0.181	0.154	0.136
FRA	level	4.654	2.399	1.633	1.248	1.019	0.866
	trend	0.829	0.452	0.319	0.252	0.213	0.187
GER	level	4.291	2.240	1.544	1.195	0.986	0.848
	trend	0.513	0.285	0.208	0.169	0.147	0.133
IRE	level	3.499	1.825	1.262	0.981	0.813	0.700
	trend	1.041	0.550	0.385	0.303	0.253	0.220
ITA	level	4.893	2.513	1.712	1.310	1.069	0.909
	trend	0.714	0.400	0.292	0.237	0.204	0.183
JAP	level	4.228	2.172	1.486	1.145	0.942	0.809
	trend	0.574	0.309	0.220	0.176	0.150	0.134
NEL	level	1.387	0.730	0.506	0.395	0.330	0.288
	trend	1.003	0.524	0.361	0.280	0.232	0.201
NEW	level	3.236	1.711	1.191	0.929	0.773	0.668
	trend	0.624	0.353	0.256	0.208	0.180	0.162
NOR	level	3.139	1.666	1.169	0.921	0.774	0.679
	trend	0.730	0.391	0.278	0.222	0.189	0.169
POR	level	3.705	1.938	1.339	1.037	0.856	0.736
	trend	0.490	0.314	0.250	0.211	0.187	0.172
SPA	level	4.414	2.285	1.566	1.209	0.994	0.851
	trend	0.838	0.453	0.317	0.250	0.211	0.184
SWE	level	3.972	2.058	1.414	1.091	0.897	0.768
	trend	1.205	0.631	0.438	0.343	0.285	0.247
SWZ	level	4.526	2.334	1.598	1.230	1.010	0.864
	trend	0.482	0.261	0.187	0.151	0.131	0.118
UK	level	3.062	1.647	1.177	0.947	0.811	0.722
	trend	0.600	0.333	0.248	0.209	0.188	0.175
USA	level	3.092	1.671	1.201	0.966	0.823	0.727
	trend	0.492	0.280	0.215	0.185	0.168	0.157

Table 11: KPSS tests for stationarity of tax wedge series

Tests			**Unemployment Rates**	**Tax Wedge**
H_0: All panels contain unit roots H_1: At least one panel is stationary				
ADF-Fisher	Trend	Chi-square	25,52	26,25
# lags: 2		Prob.	0,96	0,95
	Fixed	Chi-square	31,75	68,23
		Prob.	0,82	0,00
PP-Fisher	Trend	Chi-square	24,65	24,47
# lags: 3		Prob.	0,97	0,97
	Fixed	Chi-square	28,04	69,50
		Prob.	0,92	0,00
H_0: All panels contain unit roots H_1: All panels are stationary				
Levin-Lin-Chu	Trend	t*-modified	-1,46	-3,09
		Prob.	0,07	0,00
	Fixed	t*-modified	-2,81	-5,67
		Prob.	0,00	0,00
H_0: All panels contain unit roots H_1: A fraction of the series are stationary				
Pesaran CIPS	Tend	Z[t-bar]	-1,39	-0,91
		P-value	0,08	0,18
	Fixed	Z[t-bar]	-1,21	-3,39
		P-value	0,11	0,00
H_0: All panels are stationary H_1: Some panels contain unit roots				
Hadri	Trend	Z-stat.	72,47	80,33
		Prob.	0,00	0,00
	Fixed	Z-stat.	76,77	110,03
		Prob.	0,00	0,00

The Levin-Lin-Chu and the Hadri test require a balanced panel. Hence Portugal was excluded from the sample in the assessment of unemployment rates. For the corresponding tests for the tax wedge series New Zealand, Norway, Portugal and Spain were excluded.

Table 12: Panel unit root tests

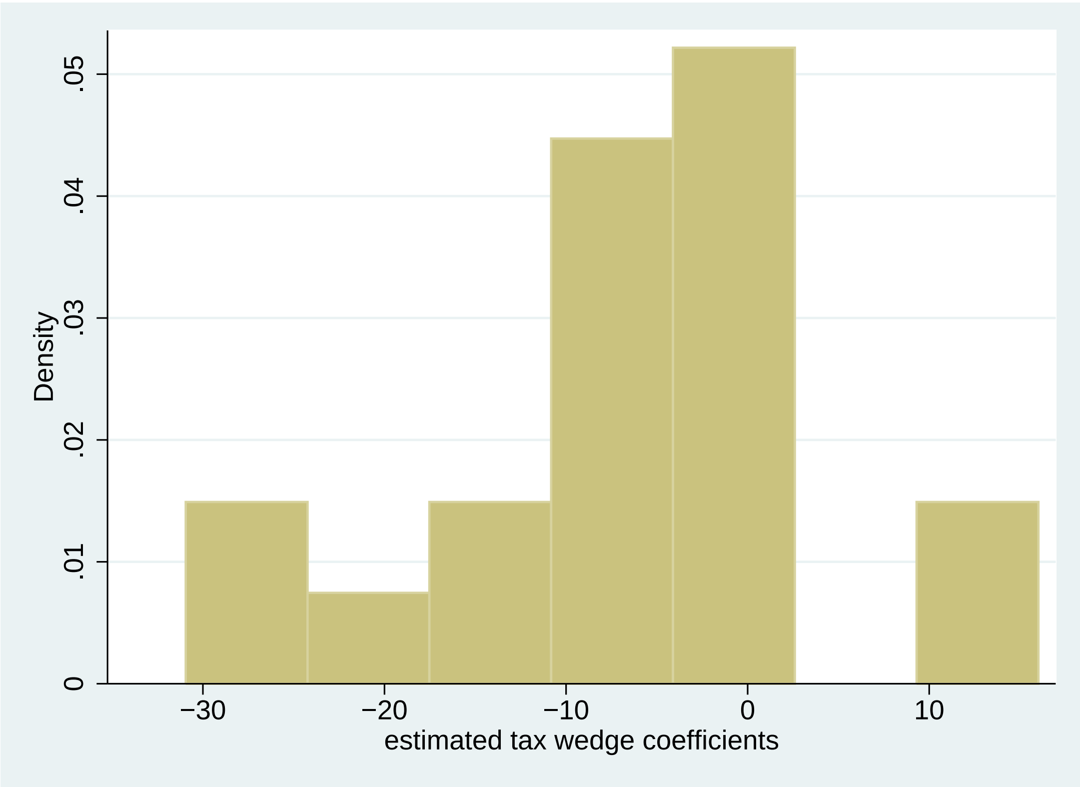

Figure 12: Distribution of SUR estimates in 'institutions-only' model

	AUS	AUT	BEL	CAN	DEN	FIN	FRA	GER	IRE	ITA	JAP	NEL	NOR	SWE	UK	USA
Δur_{-1}	0.215*	0.211*	0.444***	0.0412	-0.235*	0.341*	0.125	0.344***	0.0341	0.585***	0.437***	0.384**	-0.00860	0.150	0.398***	
	(2.02)	(2.02)	(4.46)	(0.67)	(-2.44)	(2.12)	(1.35)	(3.64)	(0.25)	(5.30)	(3.87)	(3.06)	(-0.07)	(1.21)	(3.49)	
Δur_{-2}			-0.0429	-0.118*	-0.0713	-0.253*	-0.0194	-0.295***	0.214	-0.0679		-0.370**	-0.456***	0.0416		
			(-0.46)	(-2.30)	(-0.72)	(-2.20)	(-0.19)	(-3.42)	(1.62)	(-0.66)		(-3.24)	(-4.45)	(0.34)		
Δtw	-12.82*	-5.268	18.20**	-6.437*	-7.069	6.448	6.677	-7.205	10.35	-1.123	-1.074	-1.164	-10.23**	-1.365	-9.840	7.776
	(-2.44)	(-1.54)	(2.98)	(-2.22)	(-1.13)	(1.32)	(1.22)	(-1.13)	(1.33)	(-0.88)	(-0.33)	(-0.21)	(-3.12)	(-0.47)	(-1.88)	(1.47)
Δtw_{-1}	0.438	-4.521	1.391	-0.890	-2.973	2.068	-0.601	-4.477	5.982	1.580	1.222	0.124	-0.377	1.590	11.04*	-4.659
	(0.07)	(-1.53)	(0.25)	(-0.27)	(-0.51)	(0.43)	(-0.11)	(-0.67)	(0.77)	(1.12)	(0.35)	(0.02)	(-0.12)	(0.54)	(2.16)	(-0.92)
Δtw_{-2}	0.872	-8.113**	0.176	10.44**	1.541	1.686	-5.402	-4.838	0.550	1.866	5.154	6.223	-0.981	-2.383	2.453	21.37***
	(0.18)	(-2.63)	(0.03)	(3.26)	(0.26)	(0.36)	(-1.20)	(-0.83)	(0.07)	(1.40)	(1.66)	(1.12)	(-0.32)	(-0.80)	(0.47)	(3.47)
Δbrr	0.0178	-0.0259	-0.00213	-0.0154	-0.0427	0.00617	-0.0191	-0.340***	0.132	-0.0204	0.0111	-0.0140	0.0318	-0.0542	-0.0177	-0.225***
	(0.23)	(-0.74)	(-0.04)	(-0.45)	(-0.76)	(0.15)	(-0.57)	(-3.62)	(1.05)	(-0.85)	(0.24)	(-0.34)	(1.15)	(-1.27)	(-0.20)	(-3.60)
Δtud	0.0864	-0.141*	0.147*	0.0702	0.360***	0.0544	-0.154	0.250***	-0.0228	-0.00559	0.00974	-0.279	-0.0635	-0.0337	-0.0949	-0.212**
	(1.22)	(-2.05)	(2.24)	(0.92)	(3.76)	(0.96)	(-1.50)	(3.49)	(-0.25)	(-0.12)	(0.09)	(-1.83)	(-0.87)	(-0.45)	(-1.16)	(-2.92)
TFP	-38.89***	-13.19	-40.19***	-71.98***	-49.37**	-43.86***	-48.06***	-56.70***	-52.66**	-34.67***	-6.077	-22.25***	-24.46***	-47.65***	-52.15***	-52.61***
	(-5.80)	(-0.94)	(-3.38)	(-12.68)	(-3.16)	(-3.48)	(-3.40)	(-5.39)	(-2.93)	(-3.61)	(-0.65)	(-3.38)	(-5.07)	(-4.28)	(-5.16)	(-8.49)
OUT	-11.04*	-5.051	21.87**	14.06***	4.908	9.071	2.002	10.30	20.26***	0.946	-3.728	-11.14	-10.36*	14.74**	-15.55*	5.902
	(-2.14)	(-0.97)	(3.23)	(4.46)	(0.45)	(1.82)	(0.34)	(1.86)	(3.32)	(0.18)	(-1.42)	(-1.60)	(-2.44)	(2.71)	(-2.40)	(1.74)
LDS	9.532	-7.194	-6.940	19.70**	15.33	-6.590	3.071	28.33*	12.20	20.50*	-3.423	14.83	18.39***	-9.723	37.67*	-22.11***
	(0.99)	(-0.52)	(-0.53)	(2.90)	(0.95)	(-0.65)	(0.23)	(2.05)	(0.62)	(2.03)	(-0.35)	(1.34)	(3.35)	(-1.10)	(2.36)	(-3.58)
TTS	12.90*	-5.151	-1.216	1.445	2.601	-0.390	-0.143	6.394	1.407	-10.82*	-8.334***	2.235	-1.245	1.991	10.16**	-7.593
	(2.47)	(-1.15)	(-0.52)	(0.45)	(0.37)	(-0.05)	(-0.03)	(1.18)	(0.29)	(-2.35)	(-3.53)	(0.63)	(-0.29)	(0.43)	(2.72)	(-0.82)
r	-2.108	-8.303	0.391	-10.71***	-2.060	-1.641	-0.852	5.979	0.999	-1.414	-1.926	-12.54*	-0.0181	1.448	-5.278	-10.09***
	(-1.06)	(-1.76)	(0.12)	(-5.06)	(-0.48)	(-0.49)	(-0.26)	(0.98)	(0.19)	(-0.75)	(-1.57)	(-2.20)	(-0.01)	(0.45)	(-1.79)	(-3.55)
TFP_{-1}	18.58**	7.316*									10.30*				40.59***	0.308***
	(3.06)	(2.14)									(2.09)				(3.76)	(3.73)
TTS_{-1}															-9.482*	
															(-2.25)	
r_{-1}															-2.770	
															(-1.00)	

Figure 13: Final SUR model

	(FE) Δur	(FE *xtgls*) Δur	(FE *xtregar*) Δur
Δur_{-1}	0.264***	0.0629*	0.154***
	(9.45)	(1.98)	(5.43)
Δtw	−1.589	0.0896	−1.341
	(−1.30)	(0.11)	(−1.13)
Δtw_{-1}	1.061	0.511	1.506
	(0.87)	(0.60)	(1.26)
Δtw_{-2}	2.123	2.434**	2.462*
	(1.76)	(2.95)	(2.12)
Δbrr	−0.00345	−0.0328***	−0.00640
	(−0.26)	(−3.38)	(−0.48)
Δtud	0.0579*	0.0144	0.0576*
	(2.53)	(0.86)	(2.50)
tfpsh	−27.25***	−26.93***	−36.74***
	(−11.28)	(−12.98)	(−13.49)
out	4.161**	0.896	3.648*
	(2.99)	(0.79)	(2.48)
lds	−12.92***	1.586	−3.749
	(−4.40)	(0.70)	(−1.20)
tts	0.638	−4.641***	1.221
	(0.47)	(−4.36)	(0.93)
rirl	−4.025***	−1.737*	−4.023***
	(−4.53)	(−2.39)	(−4.26)
Observations	892	656	870

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Fixed-effects models