

MPRA

Munich Personal RePEc Archive

Economic growth and welfare state: a debate of econometrics

DING, HONG

20 June 2012

Online at <https://mpra.ub.uni-muenchen.de/39685/>
MPRA Paper No. 39685, posted 27 Jun 2012 23:19 UTC

Economic Growth and Welfare State: A Debate of Econometrics

Hong Ding

Abstract

This study uses a two-way fixed effect model for panel data of all OECD nations, which includes most of the determinants of growth in previous empirical growth studies for either cross section or panel data as control variables and carefully checks possible endogeneity of the key variables of interest: welfare measures by Durbin-Wu-Hausman test. The empirical analysis shows a robust negative correlation between welfare spending rate, tax-to-GDP ratio and GDP growth. In particular, the estimates suggest that a 1% increase in welfare spending as percentage of GDP would increase the per capita GDP growth rate by 0.19%. Among three biggest components of welfare expenditure, pension spending is identified as the most important source of detrimental effect on growth while income support and public health expenditure are found to have no significant impact on growth. I also find that a 1% increase in tax revenue-to-GDP ratio would increase the per capita GDP growth rate by 0.18%. Since this estimate is close to that of welfare spending rate and welfare spending is only part of tax revenue used by government, it implies that decreasing welfare expenditure is more important and more effective for promoting growth than cutting tax. All results are subject to robustness checks including unit root test for panel data, slope poolability test, dependent variable persistence test, informal check of IV exogeneity and serial correlation test.

JEL Classification: H2, H53

The impact of welfare state on economic growth is a long-standing debate. To a great extent, this is a debate of econometrics as evidences provided by previous empirical literature are mixed and inconclusive.

Many debates of economic issues are due to endogeneity existed in the empirical studies on these topics, which has three sources: omitted explanatory variable (confounder), reverse causality and measurements-in-error. Ignoring endogeneity or using non-exogenous instrument variable (IV) to fix it can both bias the final results, with the latter probably outweighing the former on bias. The economic effect of welfare state is such an example. For decades, the economic impact of Europe's socialist (or "social democratic") welfare state has been fiercely debated among economists and politicians of left wing and right wing. If econometric studies on this impact could be based on randomized experimental data like clinical trials, such debate would never emerge. Unfortunately, facing only observational data, with very few opportunity of applying "natural experiments", the only option to get valid (consistent) estimation is to use instrumental variable (IV), at least for endogeneity test. However, incorrect use of IV estimator can worsen rather than alleviate bias, see Wooldrige (2002, p101-p104). The following section on literature review will show this problem in some previous econometric studies on the relationship between economic growth and welfare state.

The structure of this paper is as follows. Section I of Basic theory following the introduction presents some theoretical models and hypothesis on the subject. Section II is literature review for previous empirical studies on this topic, the problems of econometric specifications of them are discussed. Section III presents data description, including variable definition, data sources and time coverage of each variable; Part IV provides a preliminary analysis of the relationship between welfare and growth by graph; Part V presents the model specification, endogeneity test for welfare variables and estimation results; Part VI is a set of robustness checks; Part VII gives an evaluation of the model's performance by graphs; Part VIII presents policy simulation across countries and across time by graphs; Part IV is conclusion.

I BASIC THEORY BEHIND THE DEBATE

Generally, economic theories concerning the economic role of government has never reached consensus between economists. Because government welfare expenditure is financed by tax, the economic impact of welfare spending is equivalent to the effect of taxation on economic growth.

As Disney (2000) summarized, taxes on factors can affect the optimal level of the capital stock, although not its growth rate in equilibrium in the standard model (Blanchard and Fischer (1989)). However, in an endogenous growth model, the nature of public spending matters: expenditure on “productive” capital rather than transfers can have a positive long run impact on growth, while taxes will offset this, especially if they distort the relative returns on factors (Barro, 1990). Taxes on labor may also adversely affect the return on investing in human capital (Heckman, 1976) but the net return to human capital is itself affected by the tax structure (King and Rebelo, 1990). Taking into account that the role human capital plays in economic growth is widely accepted, welfare spending and tax collection matters for growth.

A classical theoretical model explaining the effect of taxation is Barro and Sala-i-Martin (2004, p147-148)’s extension of the Ramsey (1928) Growth model, which explicitly models consumer optimization. Using a phase diagram, they show that the imposition of taxes on the income from capital (taxes on asset income and firms’ earnings) leads to reductions in steady state capital per unit of effective labor and steady state consumption per unit of effective labor in the long run. These effects arise because the taxes reduce the incentive to save. If a government raises tax rate on the income from capital, in the course of the transition to the new steady state, both investment and consumption will decrease, leading to lower economic growth. Similarly, they illustrate the effect of government purchases financed by a lump-sum thus non-distorting tax, which is a reduction in steady state consumption per effective labor but unchanged steady state capital per effective labor. In either case, government intervention adversely affects economic growth.

There may be other channels through which welfare state affects growth. One hypothesis is welfare and tax might negatively affect people’s incentive to work or study. Although no formal theoretical model is presented for this hypothesis, Ding (2011)’s paper gives empirical evidences to support this hypothesis.

The main goal of this paper is not a survey of theory, but finding the problems causing inconsistency of the findings in previous empirical studies on this topic and using a more rigorous econometric model and more recent and comprehensive data to confirm whether data supports a significant effect of welfare on growth. If it is confirmed, I want to identify which part

of welfare expenditure is most relevant to affecting growth. I also want to derive policy implication of econometric estimation of economic impact of welfare state.

II PROBLEMS IN THE PREVIOUS EMPIRICAL STUDIES

There is a substantial literature on the relationship between welfare spending and economic growth in terms of either level or growth rate of GDP or per capita GDP. Atkinson (1995) summarized 9 such studies conducted in the 1990s, which showed mixed results. He questioned the validity of empirical evidences presented in these studies from three perspectives: possible reverse causality, lack of dynamic specification and poor measurement of the size of the welfare state. Atkinson's book (1995) argues that there is very little correlation between economic performance and welfare expenditure. Among all the studies mentioned by Atkinson's (1995) paper and book, one prominent problem is too few control variables included in the regressions for growth, so omitted variable bias is apparent. For example, Weede (1991) applies cross-sectional and pooled regression analyses of growth rates on employment in agriculture as a percentage of civilian employment, age of democracy, and social security transfers as a percentage of GDP, no convergence term or other typical conditions for convergence.

On the other side of empirical findings, Tullio (1987) finds that "the tax-financed growth in government expenditure which has occurred in the last 20-25 years has caused unemployment and slowed down the rate of economic growth during the period." Grier and Tullock (1989) find that the growth of government consumption is significantly negatively correlated with the economic growth in three of four subsamples, including the OECD. Their results are based on pooled OLS regressions on five-year averaged data for 24 OECD countries from 1951 to 1980. This averaging over time is for netting out cyclical fluctuations, but it can destroy information contained in the sample, as the authors admit. As a consequence, their OECD sample has only 144 observations. This averaging approach not only substantially decreases statistical power but also cannot "iron out" business cycles well as the starting year, ending year and interval of averaging is arbitrarily set. Arbitrary averaging across time cannot make sure the effects of booms and downturns of economy cancel out each other. In my view, the correct way of

addressing business cycle shock is including fixed effects of years in model so that global booms or downturns can be explicitly incorporated while using investment rate to capture country-specific cyclic fluctuations. The goal of empirical growth study should be investigating the effect of variable of interest on growth after controlling for these global and country-specific cyclic shocks. Averaging across time for the intervals of five years or ten years is not appropriate way to control for these two cyclic shocks. Another problem of their study is also related to the small sample size resulting from the averaging. In addition to growth of government consumption share of GDP, only four control variables are included: inflation (level and change), population growth and convergence term. The most important determinant of growth, investment/saving rate is missing. Other widely accepted growth determinants, such as human capital and international trade openness are also missing. So we have reason to believe the results of their estimation may suffer from omitted variable bias.

Barro and Sala-i-Martin(2004)'s cross section regressions of the growth rates of per capita GDP find that the estimated coefficient of the government consumption ratio (subtracting the estimated ratio to real GDP of real spending on defense and noncapital real expenditures on education) is negative and significant: -0.062 (0.023). The main weakness of Barro and Sala-i-Martin(2004)'s econometric processing, however, is that they take account of the likely endogeneity of the explanatory variables by using lagged values as instruments. The effect of government consumption, for example, may last for a period of time so that it is likely that previous government consumption affects current economic growth rate. Because the strict exogeneity of their IVs is questionable, the consistency of estimates may also be questionable.

Lindert (2003) even directly claimed that welfare state is a free lunch and argued that there is no clear net GDP cost of high tax-based social spending on GDP. However, some econometric problems in his empirical findings cast doubt on validity of his econometric estimation results. For example, he used two-stage-least-square (2SLS) estimation to test the economic effect of heavy taxation and redistribution of welfare states on growth of GDP per capita, in which the instrumental variables (IVs) he used for social spending and tax rates are age distribution, voter turnout rates, average income, religion, ethnic fractionalization, and openness to trade (table 1). It seems that we can have sufficient reasons to suspect exogeneity of these IVs for at least some

of them are almost certain to be correlated to unobservable or omitted potential determinants of per capita income growth. For example, a country with higher proportion of youth, and higher openness to international trade tends to have higher growth (see Barro and Sala-i-Martin (2004, p530) for the latter effect). Also, the effect of religion on economy is widely accepted (see for example, Barro, McCleary (2003)). Average income, of course receives feedback effect from GDP growth rate. Endogenous explanatory variable(s) may lead to bias, but if instrument variable is not truly exogenous, the IV estimate's bias is even larger than OLS estimate. In addition, he actually run two-stage least square regressions manually by replacing the values of potentially endogenous variables with the predicted values obtained from the first stage regressions on all exogenous variables. Standard econometrics tells us that this manual implementation of two-stage least square can get correct parameter estimate but incorrect standard error thus all statistical tests based on second stage t statistics may not be valid.

Beraldo et. al. (2009)'s empirical analysis based on 19 OECD countries observed between 1971 and 1998 finds that public welfare expenditure on health positively affects growth. The problem of their estimation method still is too few of growth determinants were included as control variables in addition to welfare variables. For most of their estimation results (table 1-4), only capital stock and labor were used as control variables. This kind of growth accounting method subsume all the important determinants of growth in theory (particularly convergence term and investment/saving rate) to residual of regression and inevitably brings about severe omitted variable bias, not to mention the bias from measurements-in-error arising from imprecise estimate of capital stock. They did notice the possibility of endogeneity and used IV estimation in their robustness checks part. They used lagged values (t-1 up to t-3) of the welfare expenditure variables as instruments. They also included Gini index as additional IV and checked for overidentification by using Sargan test. It is known that Sargan test for IV validity can only be used for over-identified IV based on the assumption that just-identified IV(s) is truly exogenous. Using lagged values of possibly endogenous variables as instruments is never an appropriate way to ensure strictly exogeneity of the instruments for panel data. As Angrist & Krueger (2001) pointed out, "One of the most mechanical and naive, yet common, approaches to the choice of instruments uses atheoretical and hard-to-assess assumptions about dynamic relationships to construct instruments from lagged variables in time series or panel data. The use of lagged

endogenous variables...is problematic if the equation error or omitted variables are serially correlated". As the following section VI demonstrates, growth regression always has residual error serially correlated. So Beraldo et.al. (2009)'s approach of using one to three period lagged values of endogenous variable (health spending expenditure) as instrument variables makes the exogeneity of these IVs very questionable. The strict exogeneity of Gini index is also doubtful. Growth and Gini index are highly likely to be related to each other as there might be a trade-off between economic efficiency and income re-distribution. Because of this, their conclusion that health expenditure positively affects growth may not be valid. Actually my study using one different IV for endogeneity test finds that public welfare expenditure on health has no significant effect on growth after adjusting for serial correlation.

To sum up, previous growth literature of empirical studies have the following problems: 1) too few control variables in addition to variable of interest: either no convergence term or with convergence term but too few conditions for conditional convergence, which leads to likely omitted variable bias; 2) no endogeneity test for possible omitted variable bias or bias from reverse causality; 3) using inappropriate instrument variable to correct for endogeneity bias; 4) inappropriate processing of panel data by using averaging across time.

This paper aims to specifically address these problems by using two-way fixed effect estimation for panel data of all OECD nations using the latest data from several official sources (see table 1), with instrument variable used for endogeneity test and a set of robustness checks subject to the model.

III THE DATA

This study is based on a panel model of all 34 OECD member states. The variables used in the paper, data source and time coverage of each variable are presented in table 1. Two variables are used to represent welfare state: public social welfare expenditure as percentage of GDP (public_social) and total government tax revenue as percentage of GDP (tax). The most important three components of welfare expenditure: pension benefit spending, public health expenditure and income support are also examined separately. Welfare expenditure rate is a

better measure for welfare state than government consumption as percentage of GDP because government purchases of goods and services for citizens financed by tax may have externality benefits (for example, through education and R&D) while welfare spending is more relevant to transfer payment part of government spending which is more likely to affect individual's incentive to work or individual firm's incentive to make investment so that welfare spending rate is a better measure for non-productive effect of government intervention in economy, which is the interest of this paper.

The main data source of welfare expenditure and its components is OECD Social Expenditure database for the years 1980 – 2007. Over this period, public social expenditure as a percentage of GDP, on average across OECD, increased from 15.6% to 19.2%. Public pension spending (6.4% of GDP) and public health expenditure (5.8% of GDP) are the largest social spending items (Adema et. al. (2001)).

Table 1 The variable definitions, data sources and time coverage of data

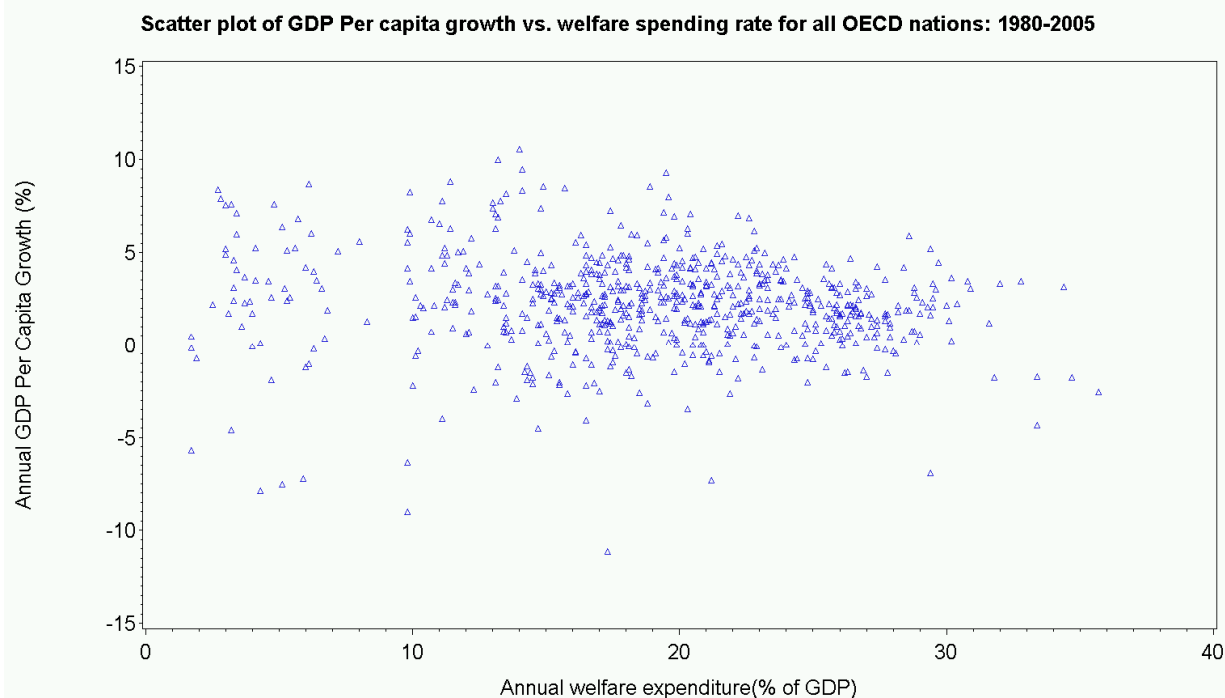
variable	Definition	Data source	Time coverage
dtot2	Change in terms of trade	WDI2010	1981 - 2010
gdppcg	GDP per capita growth rate(annual %)	WDI2010	1961 - 2005
health_exp	Public social expenditures on Health as % GDP	SOCX	1980 - 2012
hours	Average hours actually worked: Hours per year per person in employment	OECD Factbook 2010	1980 - 2012
income_support	Public social expenditures on income support to the working-age population as % GDP	SOCX	1955 - 2008
inflation	Inflation rate: Consumer price indices (CPI): annual growth in percentage	OECD Factbook 2010	1959 - 2008
invrate	Investment rate: the share of total GDP that is devoted to investment in fixed assets	OECD Factbook 2010	1976 - 2006
loglypc	Log of one-period lag of per capita GDP, measured in Purchasing Power Parity(PPP) constant 2000 international \$	WDI2010	1982 - 2009
lres	One-period lag of the number of researchers per thousand employed, full-time equivalent	OECD Factbook 2010	1980 - 2012
pension_exp	Public social expenditures on pension as % GDP	SOCX	1980 - 2012
popg	Population growth rate	OECD Factbook 2010	1951 - 2010
public_social road	Social Expenditure as percentage of GDP Road fatalities Per million inhabitants	(SOCX) OECD Factbook 2010	1980 - 2012 1990 - 2008
tax	Total government tax revenue as % of GDP	WDI2010	1955 - 2009
road_paved	The paved road as percentage of total roads	WDI2010	
trade_open	International trade openness	WDI2010	1960 - 2008

Note: WDI2010: World Development Indicator 2010 Edition, World Bank. SOCX: The OECD Social Expenditure Database

IV PRELIMINARY INFORMAL ANALYSIS

Is there any correlation between annual growth rate and annual welfare spending rate? As a preliminary check of the relationship between the two, a scatter plot of the two variables is used to illustrate the overall correlation between them for all OECD nations during the period of year 1980-2005.

Figure 1



The plot does not look to show apparent correlation between two variables, which might explain why many layman observers claim that there is no correlation between economic performance and welfare expenditure (for example, Wikipedia's web page on welfare state). On the other hand, it vividly illustrates limitation of visual inspection. It can also demonstrate omitted variable bias existed in many previous researches on this topic. This graph is essentially a regression of growth against only one explanatory variable, welfare spending rate. If such an OLS regression is run for the pooled data of all OECD nations between year 1980 and 2005, a strongly significant estimate, -0.046 , is obtained for the parameter of welfare spending rate with a p value 0.0017 . However, the validity of this estimate is highly suspicious. This naive regression does not address omitted variable bias properly so the magnitude or even direction of the estimated coefficient may be biased. Actually the result in the next section of formal analysis indicates that pooled OLS regression with no other control variables gives a much lower estimate than true

value (-0.046 vs. -0.185). The similar pattern can be found for scatter plot of growth vs. tax-to-GDP ratio, indicating the necessity to perform formal analysis.

Since our data on growth rate covers from year 1961 to 2005, the analysis of this paper is not affected by financial crisis occurred in 2008.

V THE MAIN MODEL AND RESULTS

The specification of the model used in this paper is:

$$y_{it} = \theta_t + x_{it}\gamma + w_{it}\delta + c_i + u_{it} \quad t=1,2,\dots,T \quad (1)$$

where x_{it} is 1 x 7 vector and contains 7 observable explanatory variables which are assumed to be strictly exogenous, including one state variable representing convergence term¹: lagged value of log of per capita GDP (loglypc). The rest of the explanatory variables in x_{it} are the conditions for convergence of income, which include one state variable representing human capital: lagged value of the number of researchers per thousand employed workers, five environmental and control variables: population growth rate (popg), inflation rate(inflation), international trade openness(trade_open), terms of trade shock (dtot2) and investment rate(invrate). w_{it} is a key variable of interest, one welfare measure (one of public_social, pension_exp, health_exp, income_support, otherwel, and tax), which may be endogenous. c_i is unobservable country fixed effect (FE) or country heterogeneity, which may include some institutional variables with high time constancy, such as measures of rule of law. θ_t is a fixed effect term for aggregate time, which captures global trend of some growth determinants that are common to all OECD countries, such as worldwide technology progress or global economic downturns or booms. u_{it} are idiosyncratic errors, which also absorb some omitted variables, such as financial development level or the quality of government financial regulation. This is a two-way fixed effect model for unbalanced panel data. The reason to choose fixed effect (FE) model rather than random effect (RE) model is for controlling unobservable time-invariant country heterogeneity

¹ The technical rationalization of using lagged values as convergence terms for growth for panel data is in Appendix.

and global time trend of technology despite the fact that RE estimator may have higher efficiency than FE estimator when unobservables are not correlated with included explanatory variables. The choice of seven control variables follows most of other empirical studies on growth, particularly, Barro and Sala-i-Martin (2004).

The implementation of this two-way FE model is the classical dummy variable approach: adding two sets of dummy variables for country and year, respectively to the OLS regression of (1).

V.I Test Strict Exogeneity of Welfare Variables

It is likely that changes in economic growth induce changes in welfare spending. Higher GDP growth, which is translated to higher tax income for government may increase welfare expenditure due to more resources for re-allocation. Beraldo et. al. (2009) also point out that a well documented stylized fact is that (total) expenditure in health rises with per capita GDP. On the other hand, opposite effect may arise through another channel: higher growth indicating good economy, more job opportunity and higher income for working people may be a disincentive for dependence on welfare benefits, particularly unemployment benefits. In short, there may exist reverse causality or feedback effect from growth rate to welfare expenditure, which violates strict exogeneity assumption for the latter. If this assumption fails, the consistency of FE estimate on welfare variable is questionable. Omitted variables, whose data is unavailable or unobservable to us, may also be the source of endogeneity, as discussed before.

I apply classical Durbin–Wu–Hausman (DWH) test to check whether welfare spending is endogenous in our growth regression or whether IV estimation is necessary. Davidson and MacKinnon (1993) suggest an augmented regression test, which can easily be formed by including the residuals of each endogenous right-hand side variable, as a function of all exogenous variables and instrument variable(s), in a regression of the original model. The key requirement for this approach is that we can correctly identify all other strictly exogenous variables except suspicious endogenous variable(s) and we can find a valid IV. We assume that all explanatory variables in x_{it} of (1) are strictly exogenous and we suspect that welfare measure in w_{it} may be endogenous. The IV has to be strongly correlated to w_{it} but has no direct impact on y (is uncorrelated with the unobservable error u_{it}).

The choice of IV is the trickiest part of DWH test or IV regression. Beraldo et. al. used lagged values (up to three period) of possibly endogenous variables (health spending variables). These IVs are of course strongly correlated with endogenous variables but the exogeneity of them is highly suspicious as explained in section II.

The instrument variable (IV) I choose for welfare variables is road fatalities per million inhabitants (road) whose data comes from OECD Factbook 2010. Road fatality means any person killed immediately or dying within 30 days as a result of a road injury accident. Suicides involving the use of a road motor vehicle are excluded. The justification of the validity of this IV is elaborated as follows.

Death rate from road accidents presumably cannot affect growth rate and seems to have nothing to do with the omitted variables that affect growth rate, such as financial development indicator or the quality of financial regulation. However, this rate may be related to welfare spending rate in this way: in welfare states with higher welfare expenditure by government and more generous welfare benefit programs, people tend to have more leisure time and slower life pace. Our data and other OECD data (such as OECD StatExtracts) clearly show that average annual hours actually worked per worker in EU welfare states are much shorter than their counterparts in the US whose welfare level is relatively lower. The life pace is presumably closely related to the probability of traffic accidents. To prove this, a simple fixed effect model of hours on welfare spending rate is run where hours, as defined before, is average hours actually worked per year per person in employment. The estimate shown on the row of “hours” in table 2 indicates that on average, in an OECD country, one percentage increase in welfare spending rate leads to a reduction of working hours by about four and half hour. As the second step of the test on my hypothesis of the relationship between road fatality rate and welfare level, a simple fixed effect model of road fatality rate on hours is run, the estimate shown in table 2 is 0.0576, meaning that on average, one more extra working hour increases the road fatality rate (per million people) by about 0.06. Consequently, when a simple FE regression of road fatality on welfare is run, welfare has a strongly significant estimate -2.01, implying that one percentage point increase in welfare spending relative to GDP is translated into a drop of road fatality rate by 2.01 percentage point (per million inhabitants). As the table 2 shows, similar relationships can be found for other welfare variables (income_support, pension_exp, health_exp).

Table 2 The relationships between working hours, road fatality and welfare variables

Dependent variable	Explanatory variable	coefficient	SE	P value for SE	obs
hours	Public_social	-4.4143	1.1136	0.0000	698
road	hours	0.0576	0.0176	0.001	574
road	Public_social	-2.0141	0.4415	0.0000	564
hours	Income_support	-4.0495	1.8387	0.028	710
	Pension_exp	-9.7614	2.3246	0.000	693
	Health_exp	-22.8939	3.6359	0.000	706

Note: SE=Standard Error. Obs=observation number. All the regressions include both time and country fixed effects.

Table 3 Results of IV relevance and strength test and DWH test

var	fvalue	probf	DWH_pvalue
public_social	10.53	0.0013	0.6199
income_support	5.24	0.0228	0.6576
pension_exp	6.86	0.0093	0.8089
health_exp	20.98	<.0001	0.8945
otherwelf	1.52	0.218	0.6998
tax	6.03	0.0147	0.5722

The overall IV relevance test is performed by running a FE model of a welfare variable on the IV, five exogenous environmental and control variables and two pre-determined state variables, that is all variables in x_{it} of (1). The table 3 indicates that the IV is strong IV only for public_social and health_exp. According to Sotck and Yogo (2005)'s thumb rule of F value exceeding 10 for one endogenous variable, it is weak IV for other welfare variables, except otherwelf, which is not significant, implying that road is not relevant to other welfare spending. Although road is a weak IV for income_support, pension_exp and tax, it can still be used in

Durbin-Wu-Hausman endogeneity test for these three variables for two reasons. Firstly, the IV strength test for these three variables are all significant at 5% level, actually one at 1% (pension_exp), one marginally at 1%(tax) and one at 3% level(income_support). Secondly, Because we have only one endogenous variable, the IV road is just-identified. Just-identified IV is approximately median-unbiased even when it is weak.

The Durbin–Wu–Hausman test can be performed as follows: we first regress a welfare measure on all the explanatory variables in x_{it} of (1) (loglypc, lres, trade_open, popg, inflation, dtot2 and invrate), an instrument variable for welfare (road), dummy variables for each country and dummy variables for each year and obtain the residual, \hat{v}_2 . Then we simply include \hat{v}_2 along with unity, all the variables in x_{it} of (1) and dummy variables for nations and years in an OLS regression of \hat{v}_2 and obtain the t statistic on \hat{v}_2 . The p values for the estimated parameters of \hat{v}_2 for all welfare measures are presented on the last column of table 3. We cannot find evidence of endogeneity for any welfare variable at 10 percent significance level against a two-sided alternative, so 2SLS estimation is not necessary for ensuring consistency of the estimate of welfare variable (assuming that we trust the instrument).

V.II The OLS Estimation Results

Table 4 The OLS Estimation results for welfare variables

predictor	Estimate	stde	pvalue	robust_stde	robust_pvalue	nobs	adj_Rsq
Public_social	-0.1853	0.0649	0.0046***	0.0794	0.0265**	407	0.4841
Income_support	-0.0130	0.1008	0.8973	0.0955	0.8924	409	0.4883
Pension_exp	-0.6030	0.1422	<.0001***	0.2180	0.0096***	402	0.5091
Health_exp	-0.4977	0.2313	0.0321**	0.3238	0.1348	407	0.4789
otherwelf	-0.2343	0.2070	0.2586	0.2210	0.2977	392	0.5419
tax	-0.1827	0.0547	0.0009***	0.0800	0.0299**	410	0.4397

*** significant at 1% level, ** significant at 5% level. All the regressions include explanatory variables loglypc, lres, trade_open, popg, inflation, dtot2 , invrate and dummy variables of years and countries. Stde=standard error, adj_Rsq=adjusted R square

We can see that public_socail, pension_exp and tax are significant at 1% level, which indicates that overall welfare state, in terms of both welfare expenditure and tax revenue, has negative impact on economic growth and the very part of welfare expenditure that is most relevant to detrimental impact on growth is pension expenditure. Interestingly, the magnitudes of the coefficients of welfare spending rate and tax-to-GDP ratio are quite close although tax-to-GDP ratio is much higher than welfare spending. After all, welfare spending comes from part of tax revenue for government. This suggests that to promote growth, decreasing welfare expenditure is more important and more effective than cutting tax despite both having positive effect on growth.

All the control variables are strongly significant (at either 1% or 5% level), except dtot2 (terms of trade shock), and have expected sign. The convergence term has a negative coefficient and that for human capital has a positive coefficient. International openness and investment rate have positive effect on growth while inflation rate and population growth rate have negative impacts. This finding is consistent across all six welfare measures.

VI ROBUSTNESS CHECKS

VI.I Unit Root Test for Panel Data

Because my data is a panel data set, I perform unit root test for panel data to avoid spurious regression. There are a variety of tests for unit roots (or stationarity) in panel datasets, this paper uses Fisher-type (Choi (2001)) test because it can be applied to unbalanced panel data. This test combines the p-values from N independent unit root tests, as developed by Maddala and Wu (1999). Based on the p-values of individual unit root tests, Fisher's test assumes that all the

panels contain a unit root thus are non-stationary under the null hypothesis against the alternative that at least one panel is stationary. The p values² associated with the Inverse chi-squared p statistics for the tests of unit root for all the variables used in this paper indicate stationarity of all of them. The detailed results are available upon request.

VI.II Test Slope Poolability of Model (1)

The main advantage of pooled cross-country time-series data for the analysis of growth equations is that the country-specific effects can be controlled for by using a fixed-effect estimator. However, this estimator generally imposes homogeneity of all slope coefficients, allowing only the intercepts to vary across countries.

In order to test the slope homogeneity across panel units for model (1), I change the specification of model (1) to (1')

$$y_{it} = \theta_i + x_{it}\gamma_i + w_{it}\delta_i + c_i + u_{it} \quad (1')$$

which is estimated by adding a set of interaction terms to (1), each of which is the interaction of each of 8 explanatory variable with dummy variables for 34 OECD countries. The equality of each slope in model (1'), γ_i and δ_i can be tested by an F statistic that all the coefficients estimated for each level of each interaction term are jointly zero. The F statistics and associated p values for the interaction terms for 13 explanatory variables (including 5 welfare variables) indicate that no interaction term is significant at 5% level so the null hypothesis of homogenous slope and heterogenous intercepts in model (1) cannot be rejected, verifying the validity of our model specification. The detailed results are available upon request.

VI.III Address Serial Correlation and Heteroskedasticity In FE Residuals

² All tests are ADF unit-root tests on each panel and include a nonzero drift, by using STATA command: xtunitroot fisher varname, dfuller lags(2) drift.

If u_{it} have very strong serial dependence, the usual FE standard errors obtained can be very misleading. This possibility tends to be a bigger problem with large T. Wooldrige (2001, p270) proves that if there is no serial correlation in the residuals obtained from fixed effect model for panel data, then

$$\text{corr}(\ddot{u}_{it}, \ddot{u}_{is}) = \frac{-1}{T-1}$$

where $\ddot{u}_{it} = u_{it} - \bar{u}_i$, $\bar{u}_i = \frac{\sum_{t=1}^T u_{it}}{T}$

So we can just use any two time periods (say, the last two), to test the above expression by using a simple regression. That is, run the regression \hat{u}_{it} on \hat{u}_{it-1} , $i=1, \dots, N$ (where \hat{u}_{it} is the residual from FE model (1)) and use the coefficient of \hat{u}_{it-1} , say α , along with its standard error, to test $H_0: \alpha = -1/(T-1)$ ($= -1/2$ if only the last two observations for each panel are chosen for the test regression). If rejected, there is serial correlation in FE residuals.

Using this method, it is easy to verify that all six OLS regressions in table 4 have residuals with strongly significant serial correlation. To address this serial correlation and possible presence of heteroskedasticity in u_{it} of (1), I also used robust variance matrix estimator that is robust to arbitrary form of heteroskedasticity or serial correlation (Wooldrige (2001, p275)).

The robust variance matrix estimator of $\hat{\beta}_{FE}$ is

$$A \text{var}(\hat{\beta}_{FE}) = (\ddot{X}' \ddot{X})^{-1} \left(\sum_{i=1}^N \ddot{X}_i' \hat{u}_i \hat{u}_i' \ddot{X}_i \right) (\ddot{X}' \ddot{X})^{-1} \quad (2)$$

where $\hat{u}_i \equiv \ddot{y}_i - \ddot{x}_i' \hat{\beta}_{FE}$, $\ddot{y}_{it} = y_{it} - \bar{y}_i$, $\ddot{x}_{it} = x_{it} - \bar{x}_i$, $\bar{x}_i = \frac{\sum_{t=1}^T x_{it}}{T}$, $\bar{y}_i = \frac{\sum_{t=1}^T y_{it}}{T}$

The column 5 and 6 of table 4 show standard errors computed by (2) and associated p values, which are slightly higher than those from non-robust estimator for most variables, as expected. The significance of the variables of interest, welfare variables, remains unchanged except that for

health_exp, which changes from significant at 5% level to non-significant at 10% level. After adjusting serial correlation, we do not find significant impact of public health welfare expenditure on growth.

VI.IV Test Persistence of Growth: Dynamic Specification

As a convergence (catch-up) term, lagged value of per capita income plays a critical role in the determination of income growth. It is likely that output growth has persistence so we may need to allow for a vector of lagged values of the dependent variable (gdppcg) as an explanatory variable so that growth persistence can be estimated. However, the consistency of the estimator for this lagged term needs to be taken into account carefully. Consider the extended dynamic FE model:

$$y_{it} = z_{it}\gamma + \rho_1 y_{i,t-1} + c_i + u_{it} \quad (3)$$

Wooldridge (2001, pp. 299) shows that in order to consistently estimate ρ_1 , the model needs to meet sequential exogeneity condition (conditional on the unobserved effect.):

$$E(y_{it} | z_{it}, y_{i,t-1}, y_{i,t-2}, \dots, y_{i0}, c_i) = E(y_{it} | z_{it}, y_{i,t-1}, c_i) = z_{it}\gamma + \rho_1 y_{i,t-1} + c_i \quad (4)$$

which is also called dynamic completeness conditional on c_i .

The hypothesis I want to test is $H_0: \rho_1 = 0$, which means that, after unobserved heterogeneity, c_i , has been controlled for (along with current and past z_{it}), $y_{i,t-1}$ does not help to predict y_{it} . When $\rho_1 \neq 0$, we say that $\{y_{it}\}$ exhibits state dependence: the current state depends on last period's state, even after controlling for c_i and $\{z_{it}, \dots, z_{i1}\}$. The condition (4) does not require that $y_{i,t+1}, \dots, y_{iT}$ be uncorrelated with u_{it} , so that feedback is allowed from y_{it} to $(y_{i,t+1}, \dots, y_{iT})$.

For the model (3), z is assumed to be strictly exogenous. The strict exogeneity of z and sequential exogeneity of lagged y implies that

$$E(u_{it} | z_i, y_{it}, y_{i,t-1}, \dots, y_{i1}, c_i) = 0 \quad (5)$$

Under condition (5), directly applying fixed effect estimation to the model (3) will lead to inconsistent FE estimator. As Wooldridge (2001, pp.301) points out, if the process $\{y_{it}\}$ has very persistent elements—which is often the case in panel data sets—the FE estimator can have substantial bias.

To test persistence (state dependence), first-differencing equation (3) gives:

$$\Delta y_{it} = \Delta z_{it} \gamma + \rho_1 \Delta y_{i,t-1} + \Delta u_{it} \quad (6)$$

Now under assumption (5),

$$E(y_{is} u_{it}) = 0, \quad s=1, 2, \dots, t$$

which implies the orthogonality conditions

$$E(y_{is} \Delta u_{it}) = 0, \quad s=1, 2, \dots, t-1$$

so at time t we can use $y_{i,t-1}^o$ as potential instruments for $\Delta y_{i,t-1}$, where

$$y_{it}^o \equiv (y_{i1}, y_{i2}, \dots, y_{it})$$

In our case, I will use $y_{i,t-2}, y_{i,t-3}$ to instrument $\Delta y_{i,t-1}$. To test for state dependence in per capita GDP growth rates, after allowing for unobserved country effects, the model is equation (6) with $y_{it} = \text{gdppcg}$ but without any other explanatory variables Δz_{it} . Further, so that we do not have to worry about correcting the standard error for possible serial correlation in Δu_{it} , I use just the differenced equations for each consecutive pairs of years from 1964 to 2005 for all countries separately.

The F statistic for joint significance of $y_{i,t-2}, y_{i,t-3}$ in the first stage regression for $\Delta y_{i,t-1}$ yields p-value of 0 and the R-squared from this IV relevance test regression is 0.75, indicating they are strong instruments. The 2SLS estimation of the first-differenced equation (6) without Δz_{it} gives

an estimate of the coefficient of $\Delta y_{i,t-1}$ for each consecutive pairs of years from 1964 to 2005 respectively, whose significance can be used to test the null hypothesis of no state dependence. Among 42 IV regressions, only year 1982, 1984 and 2002 have significant lagged difference of per capita GDP growth at 5% significance level, indicating that overall, there is no state dependence for $gdpcg$, we do not need to put the lag of it as regressor and there is no need to extend the econometric model to dynamic specification for our data.

VI.V Informal Test of Exogeneity of the IV

Standard econometrics tells us that there is no formal way of testing strict exogeneity of instrument variable used in 2SLS regression (over-identification test is based on true exogeneity of at least one IV). However, in our case, we can informally test some of the possible channels through which the IV, road may be correlated with u_{it} in (14).

One possible opposition to the strict exogeneity of the IV road is that road fatality rate may be related to road condition or quality that reflects infrastructure investment expenditure, which in turn may affect economic growth. In other words, IV is related to an omitted variable, road quality, which is correlated with infrastructure investment and is subsumed in u_{it} of (1).

To prove if road quality can directly affect growth, I use one proxy variable for road quality: the paved road as percentage of total roads in a country, the data of which comes from World Development Indicator (WDI) 2010 of World Bank. I add this variable, `road_paved` to model (1). The p values for the coefficient of this variable are 0.302 and 0.533 for normal and robust standard error, respectively. The correlation between road and `road_paved` is easily verified by running a fixed effect model of road on `road_paved` (with time and country fixed effects), the parameter and the associated p value for the coefficient of `road_paved` are -0.746 and 0.01, respectively, indicating that `road_paved` has higher correlation with road. These results clearly rule out the possibility that road quality of a nation can directly affect economic growth rate thus supporting the strict exogeneity of the IV road and the validity of ensuing DWH test.

VII EVALUATION OF THE MODEL'S PERFORMANCE

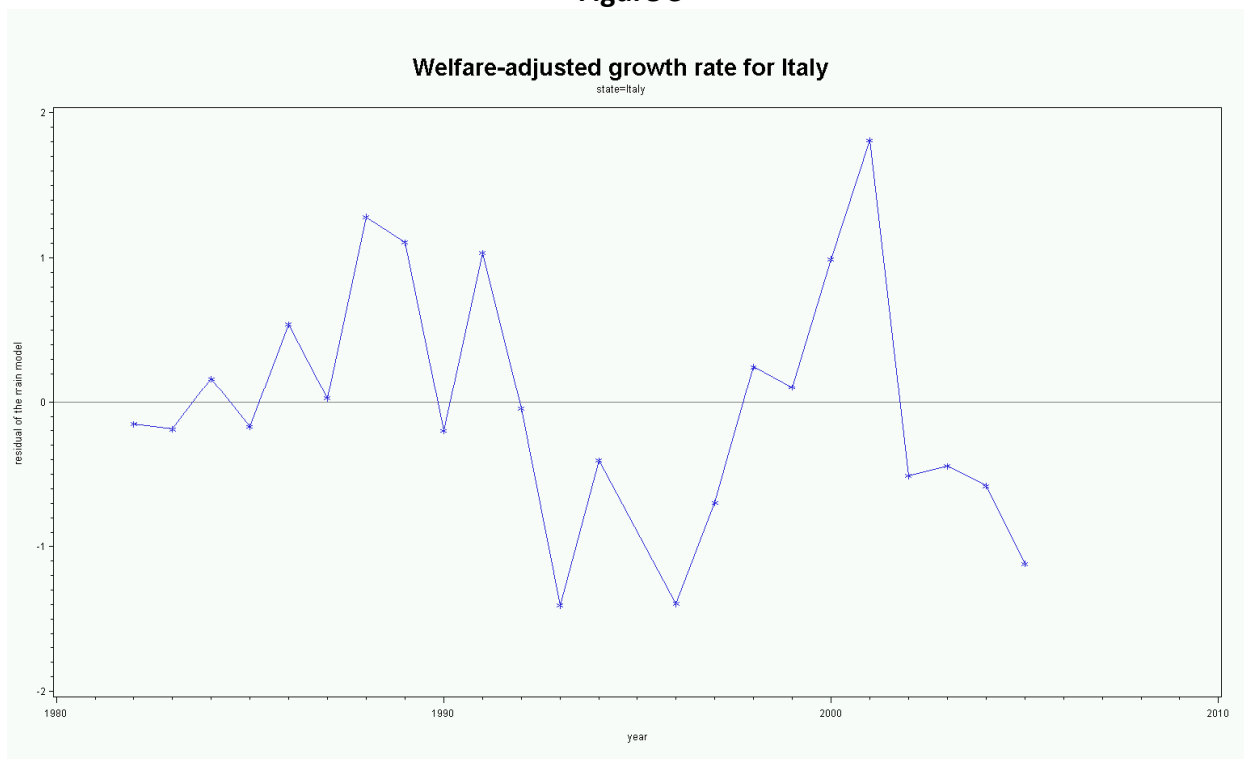
The empirical investigation of this paper seeks to answer two basic questions: first, how well do welfare variables explain the evolution of income growth over time and between countries, and second, which welfare variables turn out to matter most and why? To address the first question, Figure 2 plots some simple estimates of per capita CDP growth rates adjusted for the effects of welfare variables for USA (technically, the residuals of a panel regression of growth of model 1). While the statistical analysis confirms that welfare variables are significant determinants of growth, the chart shows that the adjusted growth rates fluctuate around zero, suggesting that cross-country differences in growth are well explained by the differences in welfare variables (public_social, pension and tax).

Figure 2



The welfare adjusted growth rate for Italy is illustrated in figure 3

Figure 3

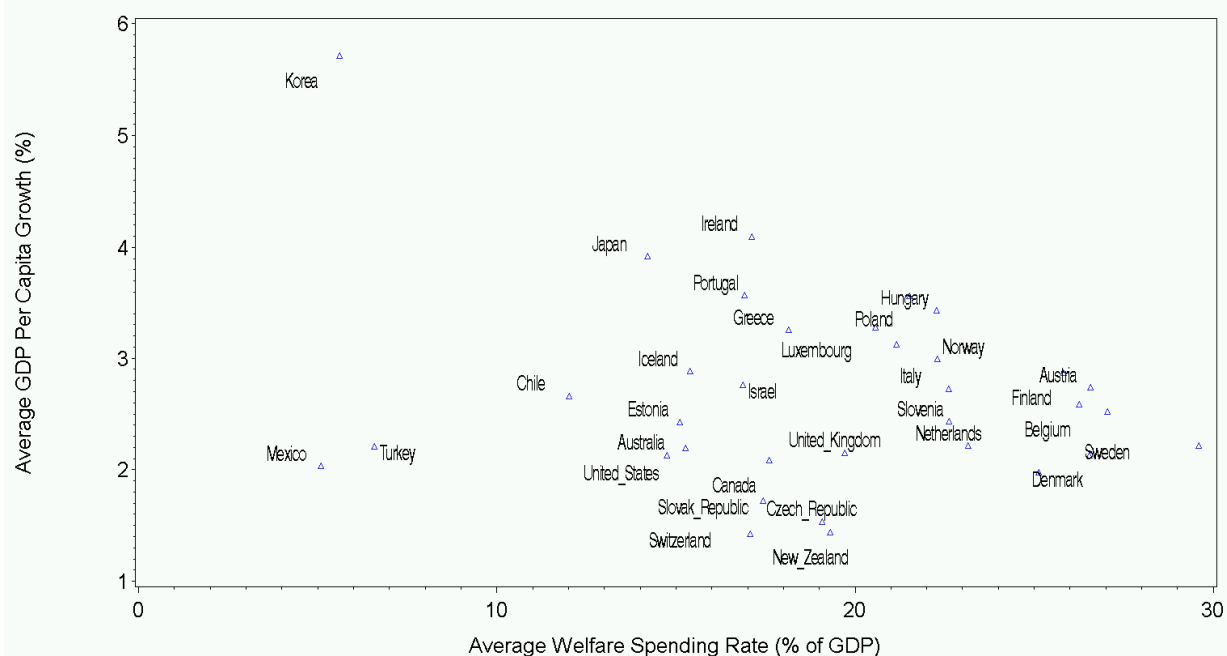


The plots of welfare-adjusted growth rates for other countries show similar pattern of no systematic trend, suggesting good explanation power and performance of our model. The plot of tax adjusted growths show similar patterns with those of welfare adjusted growths, suggesting both welfare spending rate and tax-to-GDP ratio explain growth very well.

One possible criticism of this finding may be the model is based on annual panel data thus incorporating business cycles rather than long term growth, so the significant effects for `public_social`, `pension_exp` and `tax on growth (gdppcg)` may not represent a true fundamental long-term causal relationship but a co-movements between these welfare variables and growth rate due to the shocks of business cycles (despite the inclusion of time fixed effect) impacting both of them. To address this possibility, these four variables are averaged across time for each country and a scatter plot is presented as figure 4 below.

Figure 4

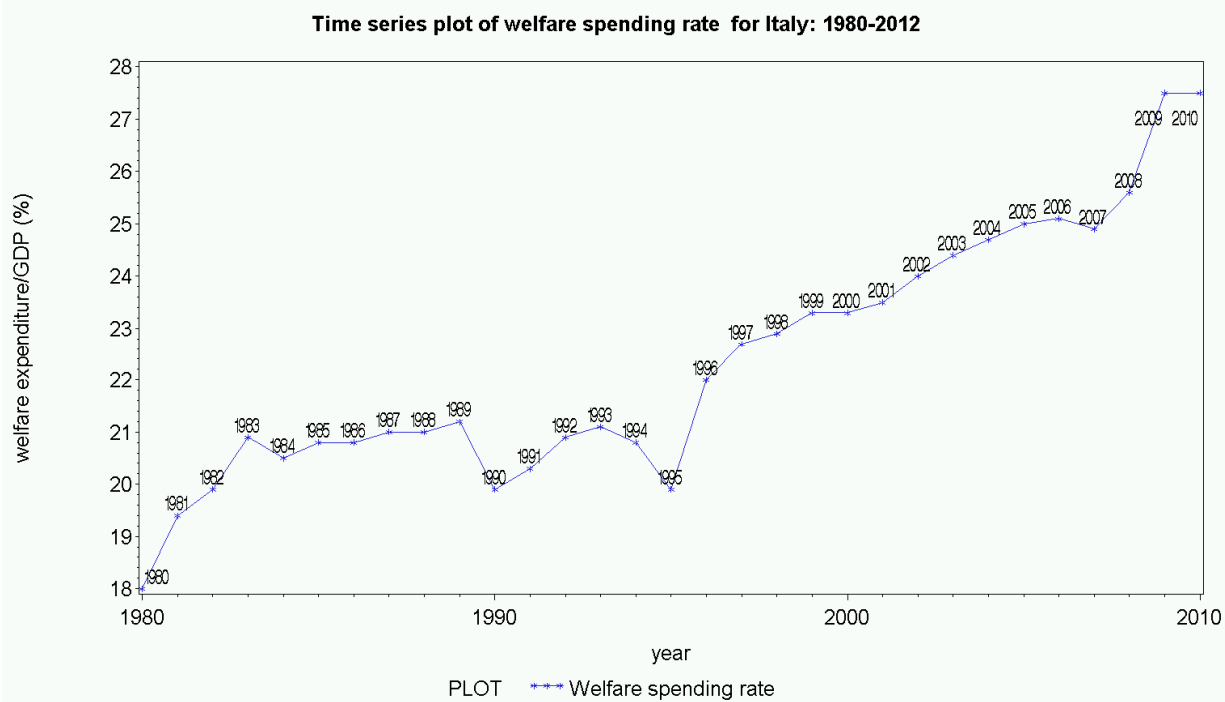
Scatter plot of average GDP Per capita growth vs. average welfare spending rate for all OECD nations: 1980-2005



We can see that except developing countries Mexico and Turkey, all other nations show an apparent negative correlation between growth and welfare spending. After “ironing” out the short-term fluctuations of business cycles, the negative correlation between growth and welfare becomes much clearer, compared with figure 1. A similar relationship can be found and illustrated between growth and tax-to-GDP ratio: except developing country Mexico and near-developed nation Chile, other nations show an obvious negative correlation. These scatter plots of course do not control for other determinants of growth, particularly convergence term, so the evidence from them is less valid and reliable than the model estimation before. However, they clearly demonstrate that the claims by many economists (Atkinson(1995), Lindert(2003)) that there is very little correlation between economic performance and welfare expenditure may not be consistent with real data.

Among 23 pairs of comparisons between actual and simulated growth rates, 17 cases show higher simulated growth than actual growth, 2 pairs overlap, only 4 cases see lower simulated growth than actual growth. Particularly, we see an increasingly wider gap between two growth rates since 1995 (with only exceptions on year 2000 and 2001). Figure 6 is the time series plot of Italy's welfare spending rate from 1980 to 2010.

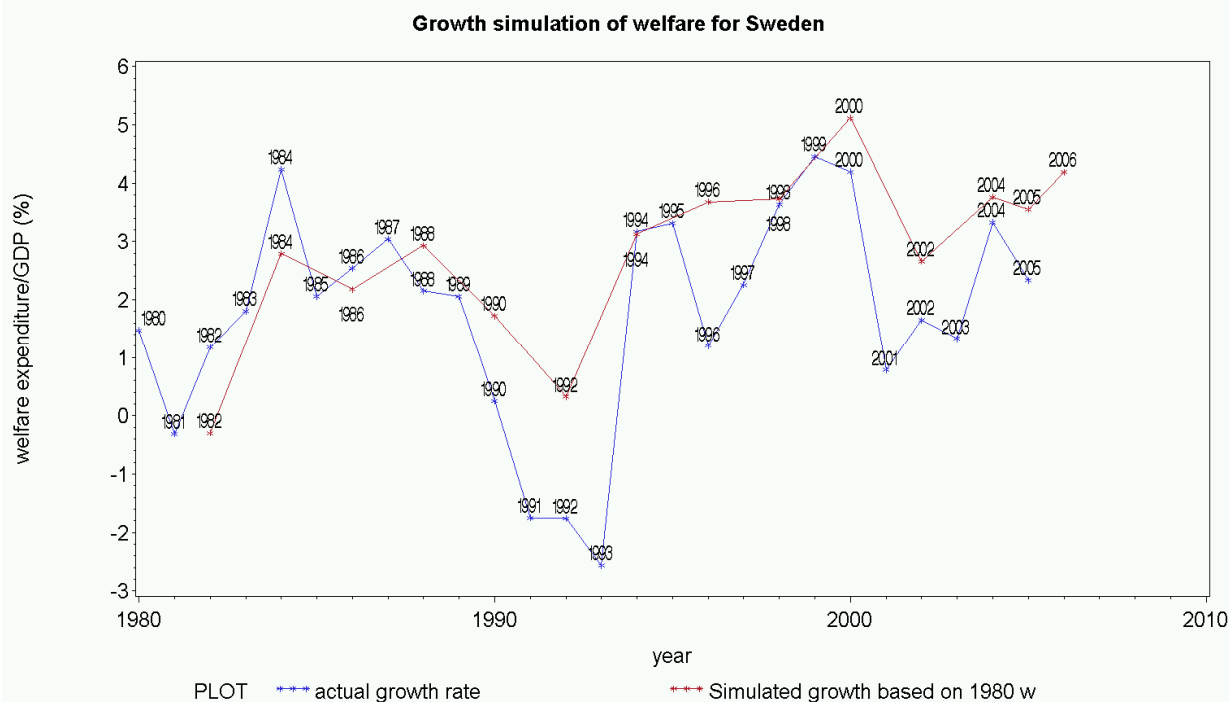
Figure 6



It is very clear that the welfare level had a great leap since 1995, which exactly match the wider gaps between actual growth and simulated growth shown in figure 5. The only exceptions on year 2000 and 2001 see little or very small change in welfare expenditure level, reflecting a nearly perfect one-to-one matching relationship between welfare spending and growth.

Having examined a country in economic crisis currently, let's take a look at the example of welfare state: Sweden. Its welfare spending rate is 27.2% in 1980, steadily increased to 35.7% by 1993 and then gradually dropped to 28.2% in 2010.

Figure 7



We can see from Figure 7 that among 13 pairs of comparisons, only three years see lower simulated than actual growth. In year 2005, if Sweden could keep its welfare spending rate at the level of 1980, its growth could have been 3.56% as compared to the actual 2.34%. Furthermore, according to our model, if Sweden had a welfare spending rate at the level of USA in that year (15.8%), its growth rate could be raised by 2.46 percentage point $((29.1-15.8)*0.1852625)$, doubling its actual growth.

IV CONCLUSION

This study uses a two-way fixed effect model for panel data of all OECD nations, which includes most of the determinants of growth in previous growth literature of empirical studies for either cross section or panel data as control variables and carefully checks possible endogeneity of the key variables of interest: welfare measures by Durbin-Wu-Hausman test. The empirical analysis shows a robust negative correlation between welfare spending rate, tax-to-GDP ratio and GDP

growth. In particular, the estimates suggest that a 1% increase in welfare spending as percentage of GDP would increase the per capita GDP growth rate by 0.19%. This estimate is close to that of Weede (1986) (-0.19 to -0.21) and slightly higher than that of Nordstrom (1992) (-0.12%). Among three biggest components of welfare expenditure, pension spending is identified as the most important source of detrimental effect on growth while income support and public health expenditure are found to have no significant impact on growth. I also find that a 1% increase in tax revenue-to-GDP ratio would increase per capita GDP growth rate by 0.18%. Since this estimate is close to that of welfare spending rate and welfare spending is only part of tax revenue used by government, it implies that decreasing welfare expenditure is more important and more effective for promoting growth than cutting tax. All these results appear to be robust after controlling for convergence conditional on human capital level, population growth, inflation, international trade openness, terms of trade shock and investment rate. As a set of further robustness checks, I also perform unit root test for panel data, slope poolability test, dependent variable persistence test (dynamic specification test), informal check of IV exogeneity and use robust standard error to correct arbitrary form of serial correlation and heteroskedasticity.

The findings in this paper supports the view that there is an inherent trend of government expansion for welfare states and without a reform aimed at downsizing welfare policies, the detrimental effect of welfare expenditures on growth would likely outweigh the positive effect of welfare programs on poverty alleviation.

Appendix

Technical rationalization of using lagged values as convergence terms for growth for panel data

In model (1) of the text, conditional convergence terms are represented by one-period lagged values of log of per capita GDP and number of researchers per million employed people. In cross section regressions for growth, convergence term is for a fixed starting year. For panel data, variable for a fixed year is unestimable. Using lagged value as convergence term is derived from Barro and Sala-i-Martin(2004)'s Log-Linearization of Ramsey model (section 2.8 Appendix 2A, p132) around the steady-state position, which can be written as

$$\log[\hat{y}(t)] = e^{-\beta t} \log[\hat{y}(0)] + (1 - e^{-\beta t}) \log(\hat{y}^*) \quad (1)$$

where $\beta > 0$, $\hat{y} = \frac{Y}{L} = \frac{Y}{L * T(t)}$ is output per unit of effective labor (the product of raw labor and the level of technology) and \hat{y}^* is the steady-state output per unit of effective labor, which is a constant.

The corresponding production function is

$$Y(t) = F[K(t), L(t) * T(t)] = F(K, \hat{L})$$

where Y is the flow of output, K is capital input (in units of commodities), L is labor input (in person-hours per year), and T (t) is the level of the technology, which is assumed to grow at the constant rate $x \geq 0$. Hence, $T(t) = e^{xt}$. Y exhibits constant returns to scale in K and L, and each input exhibits positive and diminishing marginal product.

For our panel data, the discrete version of the equation (1) can be written as

$$\log(\hat{y}_t) = e^{-\beta t} \log(\hat{y}_0) + (1 - e^{-\beta t}) \log(\hat{y}^*) \quad (1')$$

Where the subscripts refer to years, y_0 is the first year in the available panel data (in our case, year 1961).

Lagging one period for (1') and re-arrange, we get

$$\log(\hat{y}_0) = \frac{\log(\hat{y}_{t-1}) - (1 - e^{-\beta(t-1)})\log(y^*)}{e^{-\beta(t-1)}} \quad (2)$$

Substituting (2) into (1') yields

$$\log(\hat{y}_t) = e^{-\beta} \log(\hat{y}_{t-1}) + \log(y^*)(1 - e^{-\beta}) \quad (3)$$

Subtracting $\log(\hat{y}_{t-1})$ from both sides of (2) yields

$$\log(\hat{y}_t) - \log(\hat{y}_{t-1}) = (e^{-\beta} - 1)\log(\hat{y}_{t-1}) + \log(y^*)(1 - e^{-\beta}) = \alpha \log(\hat{y}_{t-1}) + c_0 \quad (4)$$

Where left-hand-side of (4) is growth rate of per capita GDP, $\alpha = e^{-\beta} - 1$ is convergence parameter and c_0 is a constant intercept.

Reference:

Adema, W., P. Fron and M. Ladaique (2011), “Is the European Welfare State Really More Expensive?: Indicators on Social Spending, 1980-2012; and a Manual to the OECD Social Expenditure Database (SOCX)”, OECD Social, Employment and Migration Working Papers, No. 124.

Angrist, Joshua D. and Krueger, Alan B (2001), Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments, *Journal of Economic Perspectives*, Volume 15, Number 4, Fall 200, Pages 69–85

Atkinson, A. B. (1995). *Incomes and the Welfare State*, Cambridge: Cambridge University Press. [ISBN 0-521-55796-8](#).

Barro, Robert J. (1990), Government spending in a simple model of endogenous growth, *Journal of Political Economy*, Vol. 98 (October), pp 103-25

Barro, Robert J.; McCleary, Rachel M. (2003), Religion and Economic Growth, *NBER Working Paper 9682*, <http://www.nber.org/papers/w9682>, May 2003

Barro and Sala-i-Martin (2004), *Economic Growth* 2nd edition, The MIT Press

Beraldo, Sergio; Montolio, Daniel; Turati Gilberto (2009), Healthy, educated and wealthy: A primer on the impact of public and private welfare expenditures on economic growth, *Journal of Socio-Economics* 38 (2009) 946-956

Blanchard, Olivier and Stanley Fischer (1989), *Lectures on Macroeconomics* (Cambridge, Massachusetts: MIT press)

Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Finance* 2001, 20: 249-272.

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

Ding, Hong (2011), Does Welfare State Affect People’s Incentive to Work or Study? An Empirical Study on Youth Inactivity Rate, in review

Disney, Richard (2000), The impact of tax and welfare policies on employment and unemployment in OECD countries, *IMF working paper*, WP00/164

Grier, K., Tullock, G. (1989), An Empirical Analysis of Cross-national Economic Growth, 1951-80, *Journal of Monetary Economics*, pp. 259-276

Headey, Bruce; Goodin, Robert E. ; Muffels, Ruud and Dirven, Henk-Jan (2000), Is There a Trade-Off Between Economic Efficiency and a Generous Welfare State? A Comparison of Best Cases of 'The Three Worlds of Welfare Capitalism', *Social Indicators Research*, Volume 50, Number 2, 115-157, 2000

Heckman, James J. (1976), A life-cycle model of earnings, learning and consumption, *Journal of Political Economy*, vol. 84 (August), pp.11-44

King, Robert G., and Sergio Rebelo (1990), Public policy and economic growth: Developing neoclassical implications, *Journal of Political Economy*, Vol. 98 (October), pp. 126-50

Lindert, Peter H (2003), "Why the Welfare State Looks Like a Free Lunch", *NBER Working Paper* No. 9869, Issued in July 2003

Maddala, G. S., and Wu, S. (1999). A Comparative Study of Unit Root Tests with Panel Data and New Simple Test, *Oxford Bulletin of Economics and Statistics*, 61, 631-652

Nordstrom, Hakan (1992), "Studies in Trade Policy and Economic Growth" Monograph No. 20, Stockholm: Institute for International Economic Studies, 1992

Ramsey, Frank (1928). "A Mathematical Theory of Saving." *Economic Journal*, 38, December, 543-559

Stock, James H. and Motohiro Yogo (2005) "Testing for Weak Instruments in IV Regression," in *Identification and Inference for Econometric Models: A Festschrift in Honor of Thomas Rothenberg*. Donald W. K. Andrews and James H. Stock, eds. Cambridge University Press, pp.80-108.

Tullio, Gissepe (1987), "Long run implications of the increase in taxation and public debt for employment and economic growth in Europe", *European Economic Review*, Vol. 31(April), pp. 741-80

Weede, Erich (1986), "Sectoral Reallocation, Distributional Coalitions and the Welfare State as Determinants of Economic Growth Rats in Industrialised Democracies" *European Journal of Political Research* 14 (1986): 501-19

Wooldridge, Jeffrey M. (2001), *Econometric Analysis of Cross Section and Panel Data*, The MIT Press, Cambridge, Massachusetts; London, England, October 2001