Taxes, Wedges, and Labor Supply: Evidence from OECD Countries, 1956-2004*

Lee Ohanian
University of California, Los Angeles

Andrea Raffo
Federal Reserve Bank of Kansas City

Richard Rogerson
Arizona State University

Preliminary

May 2006

Abstract

We document large changes in the distribution of hours worked across OECD countries over the period 1956-2004. We show that these changes translates into large and trending distortions (wedges) for the intratemporal optimality condition. We also find that most of these distortions are accounted for by changes in taxes. For few countries, instead, our calculations predict a much lower value for hours worked than observed in the data: other factors have more than offset the negative effect of taxes.

*We would like to thank Andy Atkeson, Ariel Burstein, Hal Cole, Gary Hansen, Ed Prescott and participants at the Research Department at the Federal Reserve Bank of Atlanta and Kansas City, 2006 Midwest Macroeconomic Meetings, Society for Economic Dynamics 2004 Annual Meeting and UC Davis for comments. We thank Cara McDaniel for supplying us with her tax rate series. Sonal Hate provided excellent research assistance. Rogerson thanks the NSF for financial support. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

1 Introduction

Macroeconomists have long been interested in understanding the time series behavior of hours of work. Much effort has been devoted to one particular context–changes in hours worked at business cycle frequencies in the US economy– and assessing the role of various model features in accounting for these changes. The starting point for this paper is the observation that trend changes in hours of work across OECD countries over the last 50 years exceed business cycle movements in hours of work by roughly an order of magnitude. We believe that these large trend changes warrant a systematic analysis similar to that which has occurred for business cycle movements. This paper takes a first step in this direction. In particular, we assess the ability of several versions of the neoclassical growth model to account for trend changes in hours of work for a sample of 21 OECD countries over the period 1956-2004. One goal is to isolate those episodes, i.e., time periods in specific countries, which seem most puzzling from the perspective of this theoretical framework.

Our main findings follow. First, the standard version of the growth model that abstracts from government behavior can account for virtually none of the trend decreases in hours of work observed over this period. This conclusion is robust to the value of the labor supply elasticity assumed, the inclusion of a subsistence consumption term in preferences, and to assumptions regarding how individuals value government provision of goods and services. Second, when the analysis is extended to incorporate taxes on consumption and labor income, we find that the model can account for the almost all of the average decrease in hours of work over the sample period. Additionally, the model does a much better job of accounting for changes at the individual country level. Third, although the model with distorting taxes can largely account for the average change in hours of work over time, there are several episodes that the model does not account for. Some of these episodes entail periods in which hours do not fall enough in specific countries, while others entail periods in which hours fall too much in specific countries. Importantly, those episodes which are

singled out as puzzling relative to this extension of the standard model are significantly different from those episodes which stand out as puzzling relative to the standard model. Fourth, we show in a reduced form setting that taxes are important even after controlling for a variety of other features that have been suggested as important.

Given that the one-sector growth model has become the standard framework for interpreting aggregate economic data, it is the natural starting point for trying to understand trend changes in hours across countries. To assess the ability of this framework to account for the large changes in hours worked, we employ a methodology that has been found to very useful in the business cycle literature. Specifically, we focus on the static first order condition implied by equilibrium and assess the extent to which this condition holds at each point in time in the data. One feature of this exercise is that it does not require us to assume that preferences or technology are the same across countries. We implement this exercise both for a benchmark model and for several extensions of the benchmark model, and assess how various extensions influence the extent to which the condition holds. Isolating those episodes where the condition is farthest from being satisfied serve to identify those episodes on which we need to focus and on the direction in which the model fails.

This same methodology has been profitably employed in the business cycle literature by Parkin (1988), Bencivenga (1992), Ingram, Kocherlakota, Savin (1994), Hall (1997), and Gali, Gertler and Lopez-Salido (2002). Mulligan (2002) uses this method to analyze changes in hours of work in the US over the 20th century, while Cole and Ohanian (2004) use it to shed light on changes in hours worked during the U.S. Great Depression. Chari, Kehoe and McGrattan (2002) use it to study cyclical movements of hours in two major events in U.S. history, namely the Great Depression and the 1982 recession.

The study most similar to ours is Prescott (2004). He assesses the extent to which the standard growth model with taxes can account for the changes in hours of work for a small set of countries between two particular points in time. Relative to Prescott, our contribution is fourfold, First, we extend this analysis to a much larger set of countries.

Second, we consider a much longer time period. Third, we assess the equilibrium condition for all years in our data set and not just at two distinct points in time. Fourth, we assess how several other features affect the findings.

An outline of the paper follows. In the next section we document the large and persistent changes in total hours worked across OECD countries over the last fifty years. In particular, we document the large reduction over time in the cross-sectional mean of hours worked and the large dispersion in the extent of the decrease across countries. In Section 3 we describe the benchmark model and the methodology that we employ. Section 4 presents our findings for the benchmark model, and Section 5 considers several extensions. Section 6 extends the analysis to consider distorting taxes. Section 7 carries out a reduced form analysis of the wedges from our benchmark analysis and several factors thought to be important in influencing labor market outcomes. Section 8 concludes.

2 Motivation

In this section we document three key properties of the distribution of hours worked across 21 OECD countries for the period 1956-2004:

(1) Mean hours of work have decreased substantially over the period 1956-2004.

(2) There are very large differences in the change in hours of work across countries between 1956 and 2004.

(3) There are large and persistent differences in hours of work across countries.

Our measure of aggregate hours worked is the product between total civilian employment and annual hours worked per person in employment, divided by the size of the population aged 15-64.[1]

We begin with the behavior of mean hours worked over time. Figure 1 plots the cross-country average of hours worked for each year from 1956 to 2004.

[1] Information about data sources is contained in the Appendix.

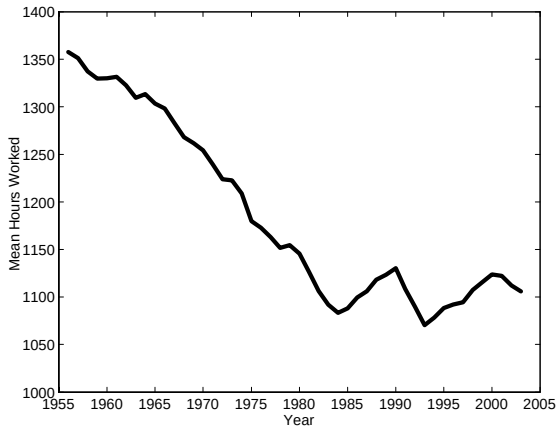


Figure 1: Mean Hours Worked

As is easily seen, this time period has witnessed a dramatic decrease in mean hours worked, with the 2004 value almost 20% less than the 1956 value. This decline occurs at a relatively steady pace over the period 1956-1985, at which point mean hours worked becomes relatively flat.

The second property that we establish is that the change in hours of work between 1956 and 2004 has been far from uniform across countries. Table 1 displays the relative value of hours worked in 2004 relative to 1956 for all 21 countries in our sample.

Table 1

Hours Worked in 2004 Relative to 1956

Australia	.97	Germany	.62	Norway	.81
Austria	.74	Greece	1.00	Portugal	.83
Belgium	.68	Ireland	.69	Spain	.86
Canada	1.07	Italy	.73	Sweden	.79
Denmark	.80	Japan	.86	Switzerland	.81
Finland	.72	Netherlands	.80	United Kingdom	.79
France	.65	New Zealand	1.04	United States	1.00

While the average change is a 20% decrease, the changes range from an increase of 7% for Canada to a decrease of almost 40% for Germany, with countries distributed throughout.

It is also of interest to examine the time series changes for each of the countries. Because we will be more interested in low frequency changes we focus on the trend component

of hours worked for each country, defined by applying the HP filter with a smoothing parameter of 100. We then divide countries into 4 groups based on differences in the trend behavior of hours worked between 1956 and 2004. Group 1 includes Austria, Belgium, Finland, France, Germany, and Italy. These countries experienced the steepest decline in hours worked in the postwar period, and exhibit a monotonic decline over time until hours worked level off around 1990. Group 2 includes Denmark, Japan, Norway, Portugal, Sweden, and the UK. Similar to group one, these countries experience a monotonic decrease in hours prior to leveling off at the end of the period, but the overall decrease is of a smaller magnitude than for Group 1.[2] Group 3 includes those countries for which there is no major trend in hours worked, namely Australia, Canada, New Zealand, and U.S. The fourth and final group consists of Greece, Ireland, Netherlands, Spain and Switzerland. A distinctive feature of this group is that the trend behavior in hours worked is not monotone. Figure A1 in the Appendix plots the HP-trend component of hours worked following the above sorting of countries into groups.

It is also instructive to consider the behavior of hours across the groups just defined. Figure 2 depicts the behavior of mean hours for Groups 1, 2 and 3.

[2] Denmark is somewhat of an exceptional case, as it exhibits a rise in hours worked over the first part of the sample period. This is due to an increase in annual hours worked per person in employment. Denmark is the only country in the sample to display an increase in this component of hours, suggesting the possibility of measurement error. In any case, from 1965 onward Denmark does display a monotonically decreasing series for hours worked.

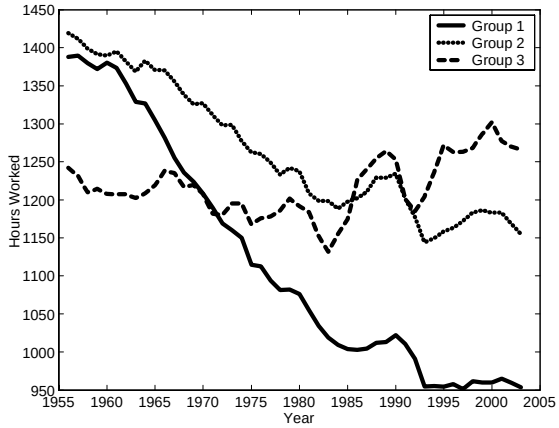


Figure 2: Mean Hours Worked by Group

Consistent with the definition of the three groups, Groups 1 and 2 both display substantial decreases over time, with Group 1 showing the larger decline, while Group 1 displays relatively little change over time.

The final property that we document is that differences in hours of work across countries are large at all points in the sample. To see this Figure 3 plots the time series for the cross-sectional standard deviation of log hours.

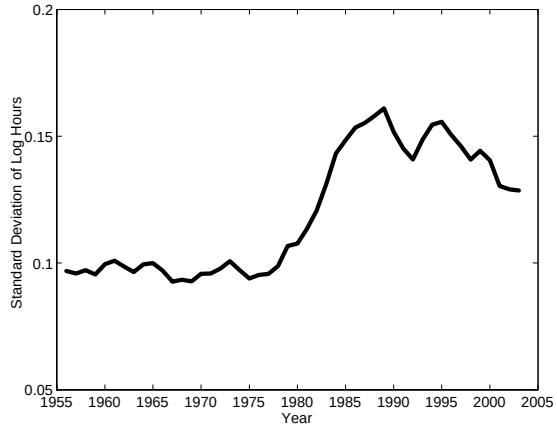


Figure 3: Disperion in Hours Worked

After a period of constant dispersion in the distribution of hours, the standard deviation of log hours worked rose steadily during the years 1975-1995, before experiencing a modest

reduction in the last ten years. While the amount of dispersion varies over time, the key point to note is that the dispersion in hours worked across countries is large at all points in time. One perspective on this magnitude is provided by noting that the standard deviation of the cyclical component of hours worked in the US over this period is .0202. It follows that the cross-country variation for this sample has typically exceeded the business cycle variation in the US by a factor of five. These differences in hours of work across countries are also very persistent over time. The average first-order autocorrelation of hours across countries exceeds .95. Recalling that our data are measured at annual frequency, this indicates a substantial amount of persistence.

To provide another perspective on the differences in hours, we take the cross-section of hours worked in 2003 and construct the distribution for the deviation of log hours worked from the mean of log hours worked. The resulting distribution is shown in Figure 4.

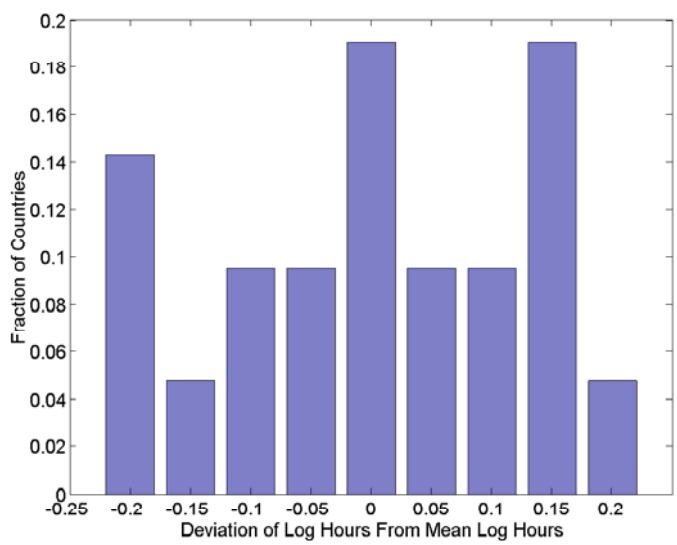


Figure 4: Dist. of Log Hours Relative to Mean in 2003

Each of the bars on the graph represents an interval of .05 centered on the value indicated beneath the bar. Figure 4 indicates a great deal of dispersion. More than half of the countries lie more than 10% away from the mean, and the gap between the mean of the lowest and highest five countries exceeds 30%. For comparison, we repeat this exercise for

the 1956 cross-section, and present the results in Figure 5.

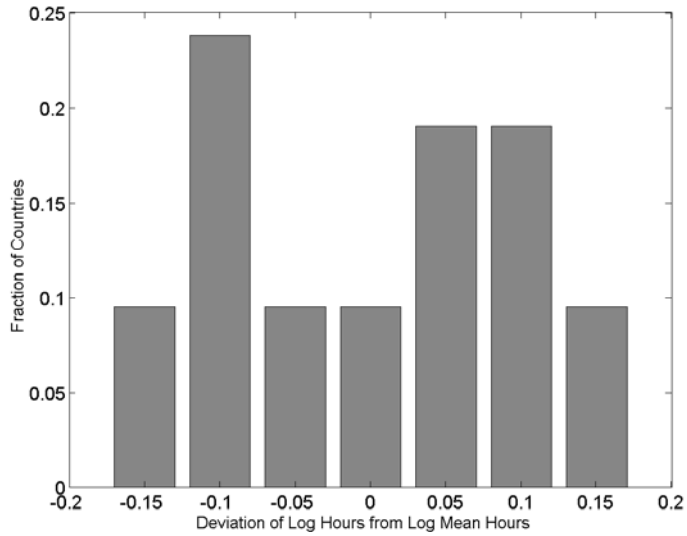


Figure 5: Dist. of Log Hours Relative to Mean in 1956

The message from this figure is the same as from Figure 4: hours of work are very dispersed in 1956.

3 Wedges in a Benchmark Model

Given that the one-sector growth model has become the standard framework for organizing and interpreting aggregate data, it is also the natural starting point for our analysis of changes in aggregate hours worked across countries. In this section we describe our benchmark model and define the concept of a wedge that serve as the focus of our analysis.

3.1 Model

While the details of this model are standard, we describe them briefly in order to introduce the notation that is needed subsequently. The economy consists of a single household with utility defined over streams of (private) consumption (C_t) and leisure ($\bar{H} - H_t$) given by:

$$\sum_{t=0}^{\infty} \beta^t U(C_t, \bar{H} - H_t) \tag{1}$$

where $0 < \beta < 1$ is the discount factor. The household is endowed with $\bar{H}$ units of time each period, and H_t represents time devoted to market work. It is standard to restrict the form of the utility function U to be consistent with balanced growth. For reasons of tractability we further restrict the utility function to be of the form:

$$U(C_t, \bar{H} - H_t) = \alpha \log C_t + (1 - \alpha)\frac{(\bar{H} - H_t)^{1-\gamma} - 1}{1 - \gamma} \tag{2}$$

where $\gamma \geq 0$, and $0 \leq \alpha \leq 1$. This form of the utility function is of particular interest because varying the parameter γ changes the elasticity of substitution between leisure and consumption, which is known to be an important quantity in analyzing how various factors influence hours of work in equilibrium.

Technology is specified by a production function:

$$Y_t = F(K_t, H_t, A_t) \tag{3}$$

where A_t is a measure of technology at time t, and K_t and H_t represent inputs of capital and labor services at time t. As is standard, we restrict the production function to be of the Cobb-Douglas form:

$$F(K_t, H_t, A_t) = A_t K_t^{\theta} H_t^{1-\theta} \tag{4}$$

Output at time t must be divided between consumption and investment (X_t) and the capital stock evolves according to the standard law of motion:

$$K_{t+1} = (1 - \delta)K_t + X_t \tag{5}$$

where $0 < \delta < 1$ is the depreciation rate.

3.2 Wedges

We focus on the competitive equilibrium allocation for this economy. Combining the first order conditions from the consumer and firm problems, one derives the so-called "static"

first order condition as one of the conditions characterizing the competitive equilibrium allocation:

$$\frac{U_2(C_t, \bar{H} - H_t)}{U_1(C_t, \bar{H} - H_t)} = F_2(K_t, H_t, A_t). \tag{6}$$

This condition asserts that in equilibrium, the marginal rate of substitution between consumption and leisure must equal the marginal product of labor at each point in time. Given the functional form assumptions above, this condition can be written:

$$\frac{(1-\alpha)}{\alpha}\frac{C_t}{(\bar{H} - H_t)^{\gamma}} = (1-\theta)\frac{Y_t}{H_t} \tag{7}$$

While this is only one of the conditions imposed by equilibrium, it has proven to be a useful diagnostic in assessing the model's ability to account for observed changes in hours, and thereby provide information as to what types of additional factors would allow the model to better account for the data. As noted in the introduction, this analysis has typically been carried out to examine the changes in hours worked within a given country at business cycle frequencies. Our goal is to use this condition to help shed light on the changes documented in the previous section.

Because our interest is in assessing the extent to which this first order condition does not hold in the data, it is convenient to introduce some notation to capture the extent to which it does not hold in the actual data. Letting H_t, Y_t, and C_t denote actual time series data for a given country, we define a time series for Δ_t for this same country by requiring that the following equation hold at each point in time:

$$\frac{(1-\alpha)}{\alpha}\frac{H_t}{(\bar{H} - H_t)^{\gamma}} = (1-\Delta_t)(1-\theta)\frac{Y_t}{C_t} \tag{8}$$

We will refer to Δ_t as the wedge in period t. Note that given our definition, a value of $\Delta_t = 0$ implies that the equilibrium first order condition holds exactly in the data. Written somewhat more compactly, Δ_t can be expressed as:

$$\Delta_t = 1 - \frac{MRS_t}{MPL_t} \tag{9}$$

where MRS_t is the marginal rate of substitution between leisure and consumption as implied by the data in period t, and MPL_t is the marginal product of labor as implied by the data in period t.

4 Results for the Benchmark Model

In this section we compute time series for the wedge for each of the 21 countries in our sample, and analyze their properties. We report results for two different calculations. In the first calculation we focus on the time series for each country considered individually. The advantage of this approach is that it does not require us to assume that preference and technology parameters are the same across countries. In the second calculation we assume that technology and preference parameters are the same across countries and compute the wedges required for the model to account for the cross-country dispersion in hours worked.

4.1 Country-Level Analysis

Letting i denote country and t denote year, define B_{it} by:

$$B_{it} = \frac{H_{it}}{(\bar{H} - H_{it})^{\gamma}} \frac{C_{it}}{Y_{it}}. \tag{10}$$

It follows directly from equation (8) that:

$$\frac{1 - \Delta_{it'}}{1 - \Delta_{it}} = \frac{B_{it'}}{B_{it}} \tag{11}$$

It is apparent that one can compute the time series values for the wedge relative to some benchmark year without any information about the values of α and θ. Given that one of our interests is in understanding what factors are able to account for the large changes in hours worked across time, it is of interest to focus on the change in the wedge over time as well, without trying to assess the absolute level of the wedge. In view of this we normalize the wedge factor to equal 0 in some benchmark year for all countries and then use equation (11) to compute the time series for the wedge. In what follows we choose 1980 as the

benchmark year for all countries. Although it is not necessary to supply values for α and θ to carry out this exercise, it is necessary to assume values for γ and $\bar{H}$ in order to compute a series for Δ_{it}. The results presented here are for the case in which preferences are log in consumption and leisure, i.e., the limiting case as γ tends to one. We discuss sensitivity of the results to the choice of γ later on. The value of $\bar{H}$ is set to $14 \times 365 = 5110$ for these calculations.

Given values for γ and $\bar{H}$, the time series for Δ can be computed given data for consumption, output and hours worked.[3] Our measure of output is GDP. For our benchmark results in this section our measure of consumption is private consumption expenditures on nondurables, services, and durables.[4] Because only the ratio of consumption to output is required, we do not deflate either series, which is equivalent to deflating both series using the GDP deflator. Using the consumption and GDP deflators to deflate the respective series does not affect our results, but it shortens the sample period[5].

The first property we report is the time series for the cross-sectional average of wedge factors across countries. This is shown in Figure 6, along with a curve showing the change in mean log hours relative to 1980.

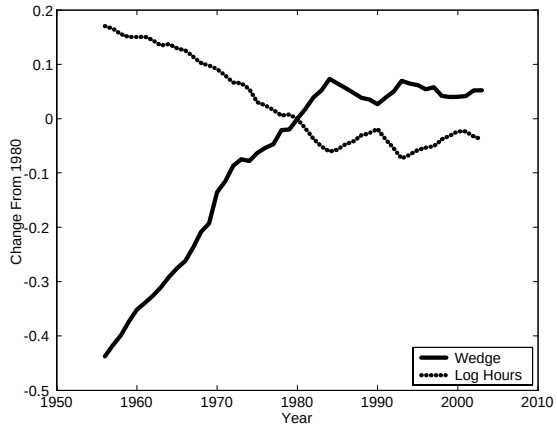


Figure 6: Mean Wedge and Mean Hours

[3] Details of the data used are provided in the Appendix.

[4] Data limitations prevent us from stripping out durable consumption goods from total consumption for a number of other countries.

[5] All these modifications of our benchmark exercise are available upon request.

The mean wedge increases at a fairly steady rate from 1956 though to the mid 1980s, at which point it levels off. Overall, the increase is about .5. Recall that given our normalization of the wedge to be zero in all countries in 1980 implies that it is only the change in the wedge that has any significance. It is of interest to compare the time series for changes in the mean wedge and mean hours worked. The negative comovement between the two series is striking, in fact, the correlation between the two series is -1.00.

Before looking at the series at the country level it is of interest to contrast the behavior of the mean wedge across the three groups defined earlier in the paper. This is done in Figure 7.

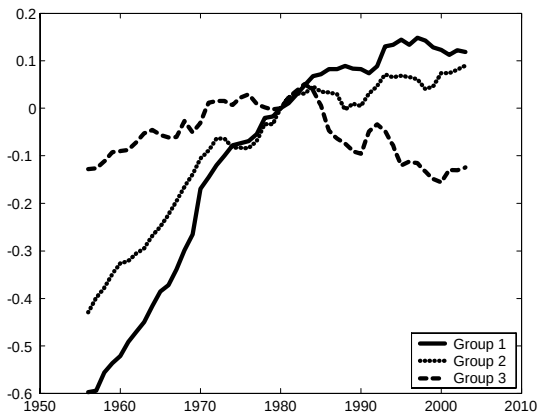


Figure 7: Mean Wedge, Groups 1, 2 and 3

Recalling the figure that showed the behavior of mean hours for these three groups, we note that there is again a striking pattern between the behavior of mean hours and mean wedges even at the group level. Group 1 exhibits a larger change (in absolute value) for both hours and the wedge as compared to Group 2. Hours and wedges are relatively flat after 1985 for both groups as well. And Group 3 displays little trend for either hours or the wedge.

We next examine the correlation between hours and wedges over time at the country level. Because our main interest is in understanding changes in the trend component of hours worked and how these changes differ across countries, rather than the business cycle

component, it is useful to distinguish between the trend and business cycle components of the hours and wedge factor series. We isolate the trend component of each series using the HP filter with a smoothing parameter of 1600. Table 2 presents the correlations between the trend components of the two series.

Table 2

Correlation of H and Δ: Trend Components

Australia	$-.86$	Germany	-1.00	Norway	$-.97$
Austria	$-.99$	Greece	$-.58$	Portugal	$-.74$
Belgium	-1.00	Ireland	$-.98$	Spain	$-.99$
Canada	$-.60$	Italy	$-.99$	Sweden	$-.99$
Denmark	$-.93$	Japan	$-.93$	Switzerland	$-.97$
Finland	$-.99$	Netherlands	$-.98$	UK	$-.97$
France	$-.99$	New Zealand	$-.81$	US	$-.98$

As the table indicates, the two series exhibit a very strong negative correlation across virtually all countries. The majority of the values exceed .97 in absolute value, and only three values are less than .80. Though we will not be interested in cyclical movements, for completeness we report the correlations between the cyclical components in Table 3.

Table 3

Correlation of H and Δ: Cyclical Components

Australia	$-.71$	Germany	$-.55$	Norway	$-.69$
Austria	$-.54$	Greece	$-.29$	Portugal	$-.46$
Belgium	$-.70$	Ireland	$-.68$	Spain	$-.66$
Canada	$-.89$	Italy	$-.48$	Sweden	$-.80$
Denmark	$-.71$	Japan	$-.24$	Switzerland	$-.82$
Finland	$-.74$	Netherlands	$-.60$	UK	$-.93$
France	$-.51$	New Zealand	$-.80$	US	$-.96$

Consistent with earlier studies, we find a very large and negative correlation for the US. The table indicates that a negative correlation is found for all the other countries, though the correlation is not as strong as for the US. The mean correlation across countries is equal to -.66.

Table 1 documented that trend changes in hours worked differed significantly across countries. In view of this it is of interest to examine how trend changes in wedges differ across countries, and how they are correlated with trend changes in hours worked. Figure 8 plots the country-level pairs for the change in log hours worked from 1956 to 2004 and the change in wedge from 1956 to 2004.

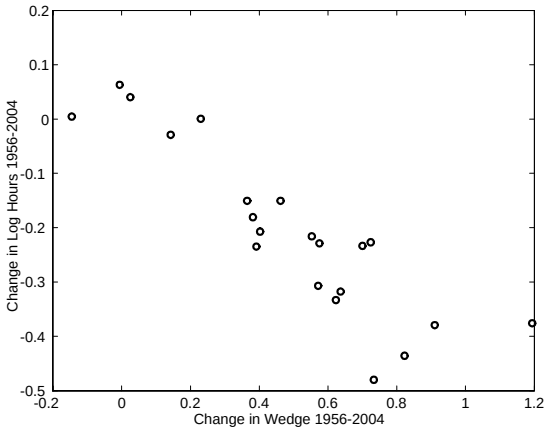


Figure 8: Change in Δ and Log H

As the figure indicates, the two are strongly negatively correlated. In fact, the correlation between the two is -.89.

While the above statistics provide a fairly strong characterization of the evolution of the wedges over time, for completeness Figure A2 in the appendix shows the trend components of the implied wedges for each country within Groups 1-3 as provided in Figure A1. For the economies in Groups 1 and 2, for which trend hours were monotonic, the wedges display similar, though inverted patterns, exhibiting a period of steady increase followed by a levelling off. For the economies in Group 3, the wedge displays relatively little trend.

The properties of changes in hours and wedges for the Group 4 countries is perhaps of greater interest. The series for trend hours in these countries display significant non-monotonicites, and it is of interest to know whether the wedges also have this property. In order to allow for a closer comparison the next four figures show the behavior of both trend hours and the wedge factors for each of these five countries.

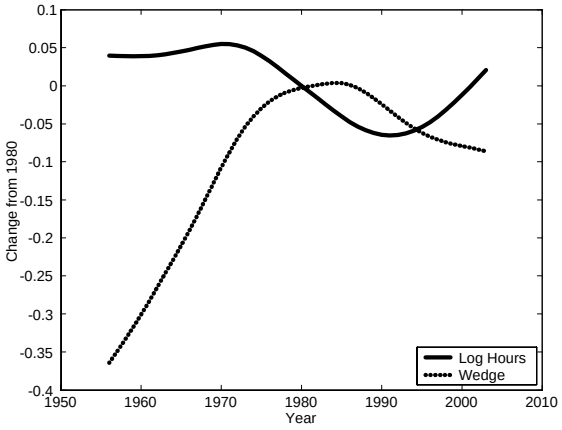


Figure 9: Wedge and Hours: Greece

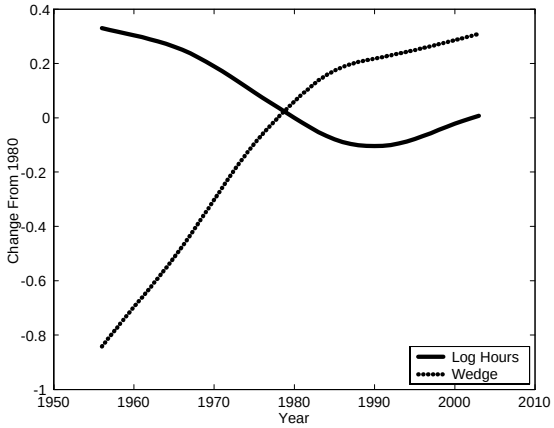


Figure 10: Wedge and Hours: Ireland

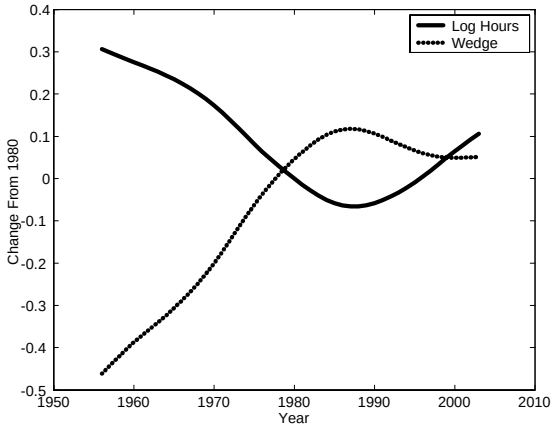


Figure 11: Wedge and Hours: Netherlands

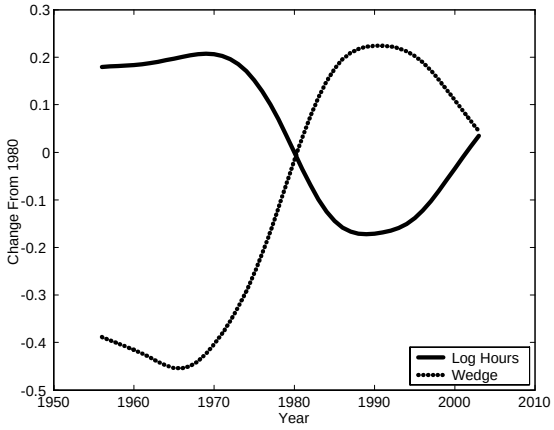


Figure 12: Wedge and Hours: Spain

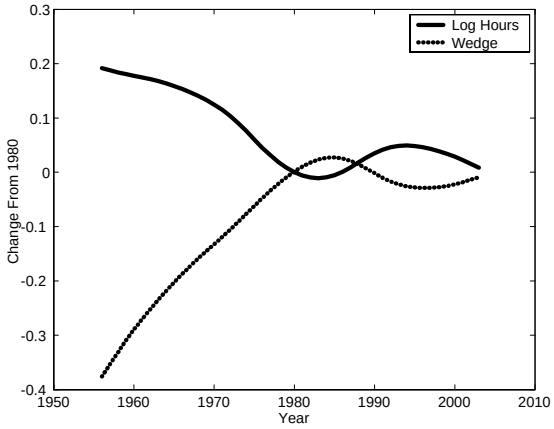


Figure 13: Wedge and Hours: Switzerland

In the Netherlands, Spain and Switzerland, the turning points in trend hours are associated with turning points in the wedge. In contrast, Ireland has a turning point in hours but no turning point in the wedge series. In Greece there are turning points in both series but the relationship is not as strong as for the first three countries just mentioned.

4.2 Cross-Country Comparison

To this point we have simply compared the changes in the wedge over time for different countries. But given the large dispersion in hours of work across countries at any point in time, it is also of interest to assess the wedges associated with the cross-section of

countries. In order to make such a calculation one must be explicit about the relative values of α and θ across countries. In this section we assume that these values are the same for all countries. In this sense our calculations should be understood as measuring the wedges that are implied by requiring that all countries have the same technology and preference parameters. Note that although the calculation does impose that θ is the same across countries, it does not impose any assumptions about the value of A_t across countries.

With all parameters constant across countries, one can treat a cross-section of observations across countries analogously to how we previously treated observations across time for a given country. In particular, we do this exercise for the 2003 cross-section and normalize the wedge factor for the US to be equal to zero. Figure 14 shows the pairs of wedges and log deviation of hours from the US value for the 21 countries in our sample.

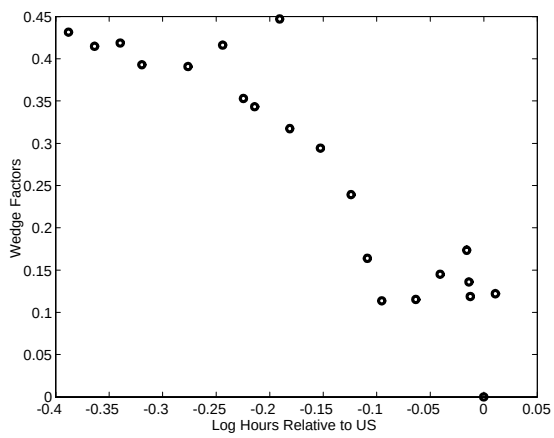


Figure 14: Log Hours and Wedge 2003

Once again, there is a striking negative correlation between the values. The correlation is equal to -.91.

This calculation can be repeated for each year of the sample. The results are summarized in Figure 15, which simply reports the correlation between hours worked and the wedge for each cross-section.

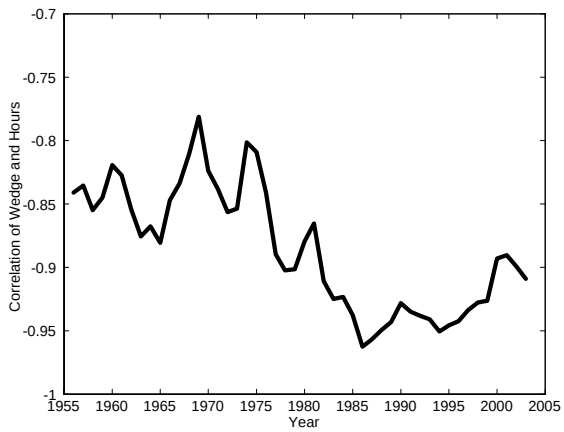


Figure 15: Correlation of H and Δ

The message is that there is nothing special about the 2003 cross-section. At all points in time, there is a strong negative correlation between hours of work and the wedge.

4.3 The Role of Consumption/Output

One simple way to describe the results just presented is that the benchmark model is able to account for neither the large changes in hours worked over time, nor the large variation in hours worked at a given point in time. The former statement holds assuming that preference and technology parameters are constant over time for all countries, though possibly different across countries, while the latter statement holds assuming that preference and technology parameters are the same across countries at any point in time. While these results tell us that there is lots of variation in hours of work that the model does not account for, it does not tell us to what extent the model is able to account for some of the variation in the data. This subsection is devoted to shedding some light on this question.

For the case of $\gamma = 1$, the expression for the static first-order condition can be rearranged to yield:

$$H_t = \frac{\frac{(1-\theta)\alpha}{1-\alpha}}{\frac{(1-\theta)\alpha}{1-\alpha} + \frac{C_t}{Y_t}} \bar{H} \tag{12}$$

This condition makes it clear that the model predicts changes in H_t only to the extent that there are changes in C_t/Y_t, with the implied changes being negatively correlated. Assessing

the extent to which differences in hours of work (either over time for a given country or across countries at a point in time) are correlated with differences in C/Y is a natural first step to answering the question laid out in the previous paragraph.

From the perspective of accounting for the change in hours worked in a given country between 1956 and 2004, the relevant calculation is the change in C/Y for that country between those two years. Table 4 shows the change in C/Y for each country in our sample over this time period.

Table 4

Change in C/Y : 1956–2004

Australia	−.05	Germany	.00	Norway	−.12
Austria	−.06	Greece	−.14	Portugal	−.18
Belgium	−.15	Ireland	−.34	Spain	−.13
Canada	−.06	Italy	−.09	Sweden	−.15
Denmark	−.21	Japan	−.08	Switzerland	−.05
Finland	−.10	Netherlands	−.12	UK	.00
France	−.11	New Zealand	−.05	US	+.08

We note two points. First, many of the changes are negative, while in order to account for a decrease in hours one would require an increase in C/Y. Second, the correlation of these changes with changes in hours worked is .29. The low correlation is evidenced by the following scatter plot for the changes in C/Y and the changes in log hours worked between 1956 and 2004.

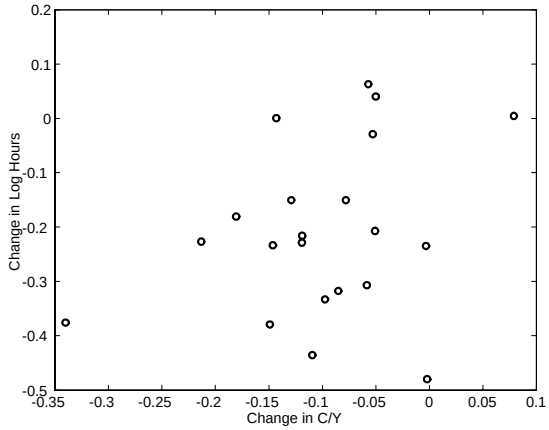


Figure 16: Changes in C/Y and H

Similar findings emerge when we examine cross-sectional differences in the value of C/Y at a given point in time. The correlation between differences in log hours and C/Y relative to the US is .36, with the scatter plot of values shown below.

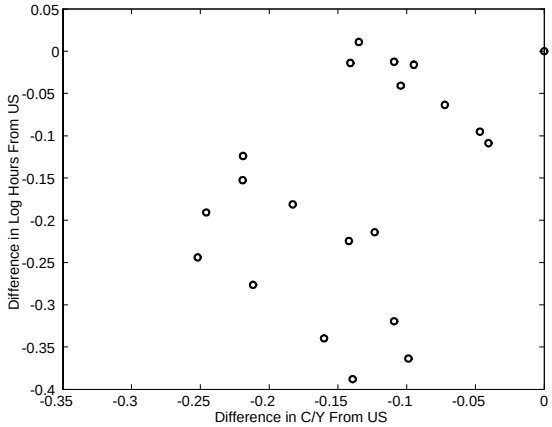


Figure 17: C/Y and H in 2003

Note that many counties have a C/Y ratio that differs substantially from that in the US, but since the difference is negative the benchmark model predicts that this should be associated with more hours of work rather than less, so that differences in C/Y actually contribute negatively to accounting for differences in hours of work. The US is also somewhat of an outlier in terms of its value of C/Y in 2004.

The preceding analysis has examined the relationship between H and C/Y at a qualitative level. The main conclusion that emerges is that from the perspective of the benchmark model, the differences in C/Y either across time for a given country or across countries at a given point in time are typically of the wrong sign in order to account for the difference in hours of work.

Nonetheless, it is also of interest to ask what the model implies for the quantitative relation between C/Y and hours of work. To assess the quantitative implications of equation (12) it is necessary to specify values for α and θ. For this calculation we assume $\alpha = .37$, $\theta = 1/3$ and an initial value of C/Y equal to .8. These values would be consistent with an investment to output ratio of .2 and time to devoted to market work equal to 1/3 of the time endowment in the steady state of the benchmark model. Changes in C/Y to either .7 or to .9 yield changes in H of 9.1% and -7.7% respectively. These results are affected very little by reasonable changes in α, θ, or C/Y.

It follows that if one wants to explain a drop of hours worked on the order of 30 percent as was observed in several countries, one would need to observe an increase in C/Y of nearly .4. And to explain differences in hours worked in the 2004 cross-section one would need the countries with low hours of work to have C/Y ratios that are as much as .5 higher than in the countries with high hours of work.

This observation also serves to indicate why it is that the results are relatively unaffected by whether consumption is deflated by the GDP deflator or the consumption deflator: the implied changes in C/Y are simply not large enough to matter a lot for the calculation of the wedge.

5 Sensitivity

In this section we examine the sensitivity of the previous findings along three dimensions. First, we ask to what extent the results are affected by the value of the labor supply elasticity parameter γ. Second, we ask to what extent the results are affected if we allow

for the possibility that government consumption is a substitute for private consumption. And lastly, we examine how the results are affected by allowing for Stone-Geary preferences that allow for a subsistence consumption term.

5.1 Sensitivity to γ

Our previous calculations were for the case of $\gamma = 1$. In this subsection we consider how our previous results are affected by this choice. The key observation is that subject to normalizing the wedge to equal zero in a particular year, different choices of γ serve simply to either increase or decrease all wedges in absolute value. While the change is not exactly proportionate, it is sufficiently close to proportionate that none of our earlier conclusions is affected. Specifically, our findings regarding relative magnitudes of changes and correlations all continue to hold.

To see the effect of changes in γ recall that Δ is defined by:

$$1 - \Delta_t = \frac{(1-\alpha)}{\alpha(1-\theta)} \frac{H_t}{(\bar{H} - H_t)^\gamma} \frac{C_t}{Y_t} \tag{13}$$

For simplicity, assume that there is no variation in C/Y. Letting $t = b$ denote the benchmark year, and following our earlier procedure that normalizes the wedge to be zero in period b, the wedge in period t is given by:

$$1 - \Delta_t = \frac{H_t}{H_b} \left(\frac{\bar{H} - H_b}{\bar{H} - H_t}\right)^\gamma \tag{14}$$

It follows that changing the value of γ changes the value of the wedge solely through its effect on the second term on the right-hand side. But this term has the property that it is strictly increasing in H_t and is equal to zero when $H_t = H_b$, independently of γ. The only effect that γ has is on the slope of the relation between H_t and Δ_t.

To illustrate the effect of changes in γ on the implied series for Δ, the next figure shows the implied time series for the trend component of Δ for France, assuming values of γ equal to 0, 1, and 2.

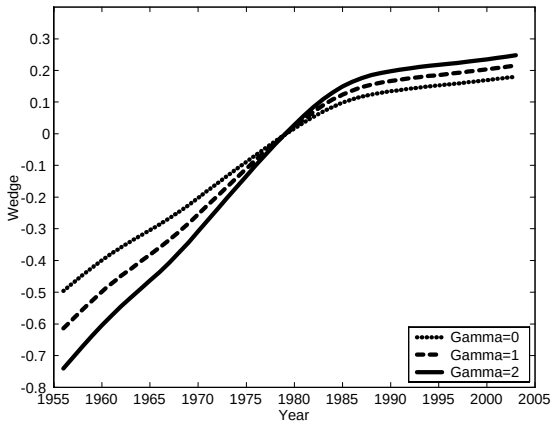


Figure 18: Effect of Gamma on Δ

As the figure indicates, in all three cases the series exhibits the same upward sloping behavior. The only effect is to alter the effect of the slope. As γ increases the increase in the wedge between 1956 and 2004 increases, being equal to .68, .83 and .99 for the cases of $\gamma = 0, 1$ and 2 respectively.

5.2 Government Consumption

The calculations that we have reported above all assumed that only private consumption entered into the household's utility function. All of the economies that we are studying have significant government expenditure on consumption, and for most countries this has increased over time. Our exercise is consistent with two different assumptions about government spending on goods and services. One is that households do not attach any value to the goods and services provided by the government. The second is that households do value these goods and services but that they enter the utility function in a manner that is separable with respect to private consumption and leisure. If government consumption is at least partially a substitute for private consumption, this could affect our findings. Governments provide many different types of goods and services, and the particular activities in which they are involved vary considerably both across time and across countries. It follows that incorporating government consumption into the analysis in the most appropriate way

is likely to be quite challenging.

However, in order to gauge the extent to which this might affect our results we can learn quite a lot by examining the extreme case in which all government expenditure on goods and services is viewed as a perfect substitute for private consumption. This amounts to simply redoing the previous calculations except that we replace C/Y with $(C+G)/Y$, where G is government expenditure on goods and services. A very simple finding emerges when we do this exercise. The effect of this change is to reduce the magnitude of the wedge by about 1/3, but to leave all of the previous properties unchanged.

This is well-illustrated by looking at the effect on the time series for the mean value of the wedge. Figure 19 shows the two series based on whether government expenditures on goods and services are included.

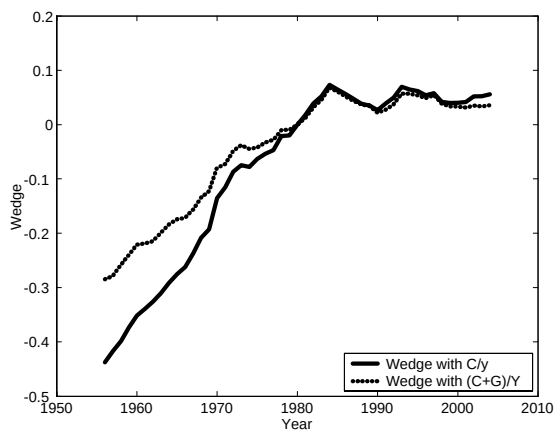


Figure 19: Effect on Δ of Including G

This figure shows that the two series move together very closely, with the big difference being that the original series exhibits a change of roughly .5 over this period, while the series based on using $(C+G)/Y$ exhibits a change of roughly .35.

This change also does not affect our results concerning the extent to which differences in C/Y can account for differences in hours of work, either across time within a country, or across countries at a point in time. The main effect is that differences in $(C+G)/Y$ are much less than differences in C/Y both in the cross-section and the time series, but the

correlations of differences in $(C+G)/Y$ and H remain weakly positive in both cases. In fact, the reason that the wedges decrease when one uses $(C+G)/Y$ is that by decreasing the differences in this ratio we are decreasing the impact of a factor that for the most part goes in the wrong direction.

One may conjecture that very different results might be obtained if the nature of government consumption changes over time and differs across countries so that the extent to which it substitutes for private consumption is changing. While such a possibility certainly increases the potential for variation in $(C+G)/Y$ to play a greater role, we report two simple calculations here to suggest that it is somewhat unlikely for this role to be a dominant one.

The first calculation is concerned with the time series changes within countries. Consider a country such as France, which exhibits one of the larger decreases in hours of work between 1956 and 2003, equal to 35%. In order to be consistent with this decrease, the model would require that C/Y increases during this period. An extreme lower bound for C/Y in 1956 for France would be the ratio of private consumption to GDP. And an extreme upper bound for C/Y in 2003 would be the ratio of private plus government consumption to GDP. But even with this extreme case the increase in C/Y is only from .67 to .80. With the values of α and θ used earlier, the predicted decrease in H is only 11%.

The second calculation carries out a similar exercise in the context of the 2003 cross-section. Again, an extreme lower bound on C/Y for the US is the ratio of private consumption to GDP, while an extreme upper bound for other countries is the ratio of private plus government consumption to GDP. This lower bound for the US is equal to .7, while the extreme upper bound is less than .8 except for three countries: Greece, the UK, and Portugal. Again, even extreme scenarios do not suggest that variation in C/Y plays a large role.

5.3 Subsistence Consumption

One of the features in the data is that almost all countries experience a decrease in hours between 1956 and 1985. There is substantial evidence to indicate that the process of development is associated with a decrease in the workweek before it seems to level off.[6] Stone-Geary preferences of the form:

$$U(C_t, \bar{H} - H_t) = \alpha \log(C_t - \bar{C}) + (1 - \alpha)\frac{(\bar{H} - H_t)^{1-\gamma} - 1}{1 - \gamma} \tag{15}$$

can account for this phenomenon. With preferences of this form, a country that experiences ongoing technological change that is labor augmenting will asymptotically converge to a balanced growth path in which hours of work are constant, but during the transition to this asymptotic balanced growth path the economy will experience a decrease in hours of work. It is only in the limit that as $\bar{C}$ becomes very small compared to C_t that the balanced growth path is achieved. In this regard it is noteworthy that the four countries that do not experience any significant drop in hours worked in the early part of the time period being studied, Australia, Canada, New Zealand and the US, are all among the most productive countries in the sample as of the beginning of the period. In view of this it is of interest to explore the possibility that the dynamics associated with subsistence consumption might be relevant in accounting for some part of the decrease in hours of work during our sample period.

Repeating the previous analysis, we obtain the following expression for the wedge:

$$1 - \Delta_t = \frac{(1 - \alpha)}{\alpha(1 - \theta)} \frac{H_t}{(\bar{H} - H_t)^{\gamma}} \frac{C_t - \bar{C}}{Y_t}$$

In order to implement this extension one needs to assume a value of $\bar{C}$. If one wants to have $\bar{C}$ constant across countries then it is also necessary to have consumption measured in common units across countries. Due to some missing observations, in order to have a complete time series for all 21 countries it is necessary to start in 1966.

[6] See for example, Maddisson (1995) for estimates of changes in the annual hours of work per person in employment.

We report results here that correspond to a value of $\bar{C}$ equal to 5% of total US consumption in 1960. For these results we assume that the measure of consumption is the sum of private consumption plus government expenditure on goods and services. The results were quite similar when $\bar{C}$ was set to 10% of total US consumption in 1960. We begin by showing the effect that this has on the mean wedge series, as compared to our original calculations. Figure 20 shows the results for the trend components.

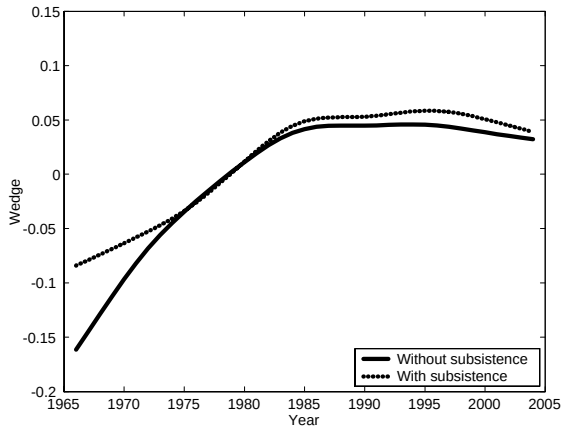


Figure 20: Effect of Subsistence on Δ

Three features are worth noting. First, the two curves have the same qualitative features. They both exhibit relatively steady increases up until the mid 1980s, are relatively flat thereafter. Second, the magnitude of the change is somewhat larger for the case without subsistence consumption. Third, in the post 1980 period, the two curves are virtually identical. This is perhaps not too surprising–since the effect of the subsistence term diminishes as consumption increases, we expect that the effect of subsistence will decrease over the course of the period.

Once again, although this extension to the benchmark model does reduce the magnitude of the wedge, all of the properties that we previously established continue to hold: differences in hours worked are associated with large differences in wedges, both in the time series and in the cross-section. As one additional illustration of this, in Figure 21 we plot

the pairs of changes in log hours and changes in wedges between 1966 and 2004 for the 21 countries in our sample.

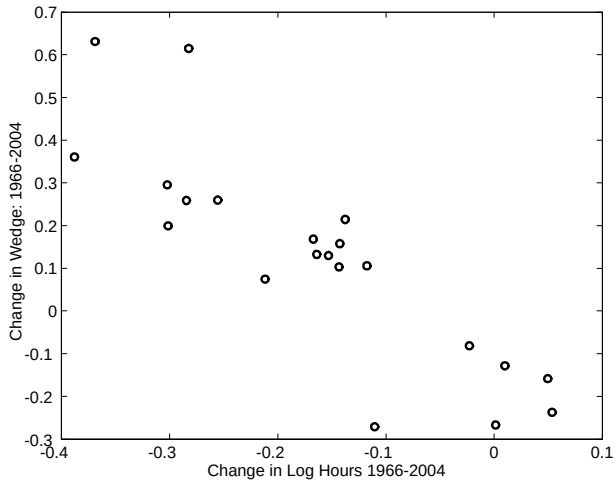


Figure 21: Change in H and Δ with Subsistence

As before, this plot shows a strong negative correlation between the two.

5.4 Summary

Having explored sensitivity along three dimensions, and presented a substantial amount of additional information, we think it is important to summarize what we view to be the main message from these exercises. Although there are variations/extensions that serve to reduce the magnitude of the wedges that we found earlier, the key message that emerges is that

6 The Role of Taxes

A simple conclusion emerges from the preceding analysis. Any extension to the standard growth model that is going to successfully account for the large changes in hours of work over time across countries must be capable of producing a "static" first order condition that in equilibrium contains a term like the wedge that we introduced earlier. It is relatively

well-known that various factors can give rise to such wedge factors, including taxation, product market regulation, non-competitive wage setting, and labor market regulation. In this section we extend the original model to include a government which taxes labor income and uses the proceeds to both purchase goods and services and fund a transfer program. Our motivation for focusing on taxes is two-fold. First, prior work by Prescott (2004) and Rogerson (2005) has suggested that taxation may be a quantitatively important factor for a small set of countries. Second, though there are issues associated with measuring tax rates, from the list just offered, taxes may be the factor which is easiest to measure, making them a natural first choice.

We carry out two exercises in this section. First, we extend our previous model to allow for a government sector, and then compute the wedge factors implied by this model. The importance of taxes in accounting for the changes in hours over time will be indicated by the extent to which the wedge factors computed for this model are reduced relative to those computed earlier. Second, we carry out a simple reduced form exercise to assess the empirical plausibility of the other factors playing a large role in accounting for the factors.

6.1 The Model with Taxes

The only change that we make to the earlier model is to add a government sector. We assume that at date t the government places a proportional tax of τ_{ht} on labor income and τ_{ct} on consumption expenditures. Government spending on goods and services in period t is given by G_t, and the remaining tax revenues are used to finance a lump sum transfer T_t so as to balance the budget each period. We modify preferences to explicitly allow for the possibility that consumers value the goods and services purchased by the government. In particular, we assume a one period utility function of the form:

$$U(C_t, \bar{H} - H_t, G_t) = \alpha \log(C_t + \lambda G_t - \bar{C}) + (1 - \alpha)\frac{(\bar{H} - H_t)^{1-\gamma} - 1}{1 - \gamma} \tag{16}$$

While the assumption that private and public consumption are additive is somewhat restrictive, as previously noted, consideration of the two extreme cases of $\lambda = 0$ and $\lambda = 1$

will be sufficient for us to bound the importance of G for the analysis.

Solving for a competitive equilibrium in this case, and following the same steps as before, one can derive the following modified version of the "static" first order condition:

$$\frac{U_2(C_t, \bar{H} - H_t, G_t)}{U_1(C_t, \bar{H} - H_t, G_t)} = \frac{(1 - \tau_{ht})}{(1 + \tau_{ct})} F_2(K_t, H_t, A_t) \tag{17}$$

To simplify notation it is convenient to define a single tax term τ_t by:

$$(1 - \tau_t) \equiv \frac{(1 - \tau_{ht})}{(1 + \tau_{ct})}. \tag{18}$$

Given our functional forms, this equation now reduces to:

$$\frac{(1 - \alpha)}{\alpha} \frac{H_t}{(\bar{H} - H_t)^\gamma} = (1 - \tau_t)(1 - \theta) \frac{Y_t}{(C_t + \lambda G_t - \bar{C})} \tag{19}$$

and if we introduce the wedge factor as before we now have:

$$\frac{(1 - \alpha)}{\alpha} \frac{H_t}{(\bar{H} - H_t)^\gamma} = (1 - \Delta_t)(1 - \tau_t)(1 - \theta) \frac{Y_t}{(C_t + \lambda G_t - \bar{C})} \tag{20}$$

6.1.1 Results

We now evaluate how the introduction of a government sector influences the size of the implied wedge factor over time across the countries in our sample. Again we consider the $\gamma = 1$ case as our benchmark. In order to compute the wedge factor we need series for both labor and consumption taxes, plus a value for λ. We use series for average taxes from McDaniel (2006), who extends the procedure of Prescott (2004) to a larger set of countries and a longer time period. We note that we do not make any adjustment to turn average tax rates into marginal tax rates. Due to data limitations we do not have series for New Zealand or Portugal and hence will only present results for 19 countries in what follows. Additionally, data was not available until 1960 for two countries, so we will further restrict the analysis to the period 1960-2003. Consistent with our earlier results, it turns out that the value of λ matters very little for most of the countries in our sample. For the results presented below we assume that $\lambda = 1$.

We begin by showing how the inclusion of taxes affects the mean series for wedges. For comparison, Figure 22 shows both this series and the comparable series derived assuming no taxes.

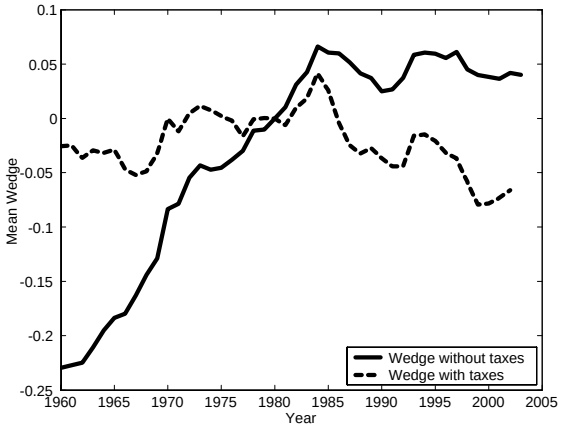


Figure 22: Mean Wedge With Taxes

The change is rather striking. Whereas the original series exhibited a large positive trend, the new series possesses virtually no trend. It follows that this model is able to account for the fact that hours of work have decreased about 20% on average.

Next we examine the patterns at a more disaggregated level. We begin by displaying the mean wedge series for Groups 1-3 as before. (Noting that Portugal and new Zealand are no longer included.) These are contained in Figure 23.

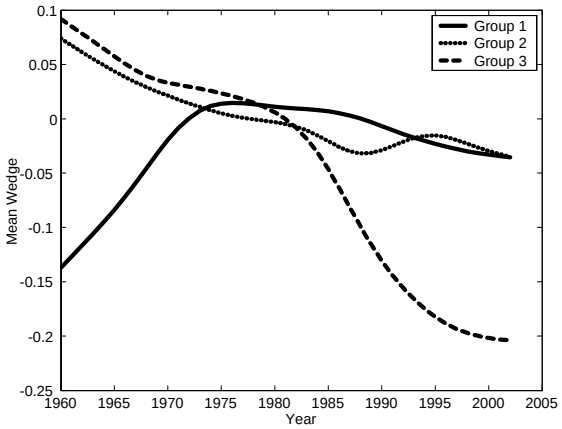


Figure 23: Mean Wedges, Groups 1-3

Several interesting patterns emerge. First, note that the series for Group 1 still displays a period of steady increase, but that this period now ends in the early 1970s as opposed to the mid 1980s. Moreover, recall that in our initial analysis the mean wedge increased by about .8 for this group, while in this figure the increase is only about .15. For Group 2, the profile has switched from having a substantial upward trend to having a relatively small downward trend. It decreases by a little more than .1, and this decrease occurs during the first 30 years. The implication is that once the model is augmented to allow for taxes, it is now unable to explain why hours in these countries did not fall more. Finally, the series for Group 3 shows a substantial downward trend, and one that accelerates around 1980. Overall the decrease is about .3. In contrast, the wedge series for this group from the original analysis was relatively flat. For this group as well, the puzzle becomes how to explain why hours worked have not fallen.

Thus far we have presented our results by reporting properties of the wedge series. Another calculation that provides a useful perspective on the ability of the model to account for the data is to solve for the series of hours of work for each country that is consistent with no wedge given observed values for C, G, and Y. Comparing this series to the actual series for hours of work provides some perspective on the extent to which the model is able to account for the data. The advantage of this approach is that it is perhaps easier to interpret a gap between two hours series than it is to interpret the significance of a wedge of a given magnitude.

Because our wedges are all normalized to be zero in a particular year, when we do the calculation for hours we also need to normalize hours in a particular year. In what follows we have normalized hours so that mean hours are the same as in the actual data. Rather than inundate the reader with figures for each of the countries, we have chosen to simply present a few that we feel are illustrative of the general results. The countries that we show are the Netherlands, Germany, Sweden and the US, with the results shown in Figures 24-27.

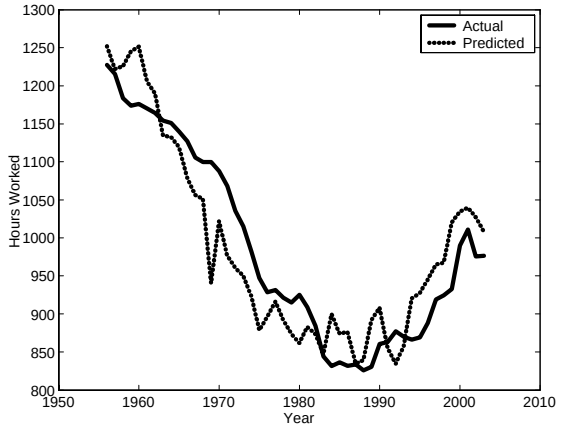


Figure 24: Actual and Predicted: Netherlands

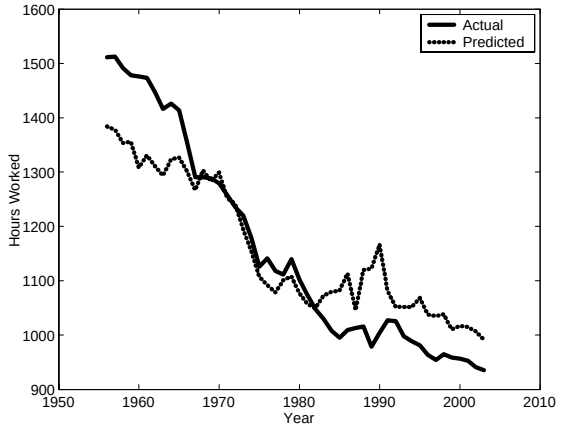
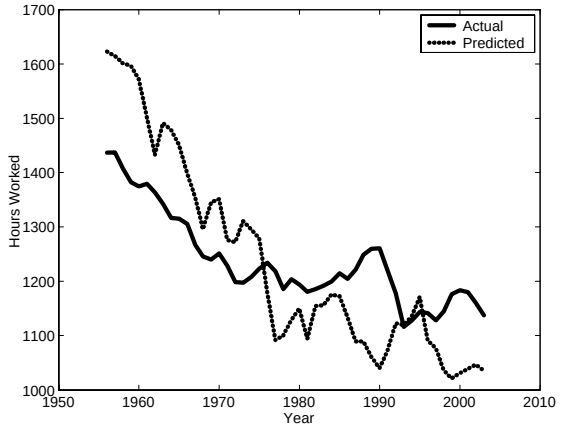


Figure 25: Actual and Predicted: Germany



Figurre 26: Actual and Predicted: Sweden

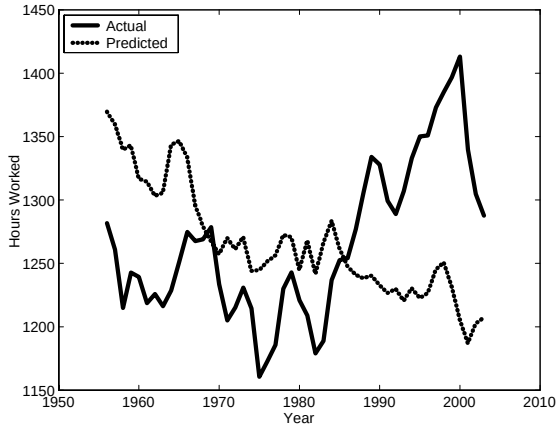


Figure 27: Actual and Predicted: US

As noted, we chose these four because they illustrate the variety of outcomes. The Netherlands is a case in which actual and predicted hours are very similar. Note in particular that the model is able to capture the non-monotonic behavior of hours. Other countries for which this is the case are Belgium, Denmark, Finland, Japan, Norway, and Austria. Germany is a case in which the model is able to account for a large share of the decline in hours worked, but not all of it. Other countries fitting this pattern are France, Italy, Switzerland, and the UK. Sweden is a case where the model actually predicts a larger drop than what we see in the data. The other country with this pattern is Greece. Finally, the US is a country which shows little trend in hours of work, but for which the model predicts that there should be a decrease. Other countries fitting this pattern are Australia and Canada.

6.2 Puzzling Episodes

One of the useful products of our analysis is a comprehensive analysis of which time periods in which countries are puzzling from the perspective of the theoretical framework that we have employed. To the extent that the theoretical framework is viewed to be a natural benchmark, this information should be very useful in helping to focus future research efforts in the area.

In an attempt to convey this information, in this section we provide decadal changes in the trend components of wedges for each country. From this one can assess which country-decade pairs are the puzzling episodes. Table 5 presents the results.

Table 5

Changes in Wedges by Decade

	60-70	70-80	80-90	90-00	60-00
Australia	−.04	−.07	−.05	−.05	−.21
Austria	.04	.00	−.03	.04	.05
Belgium	.10	.02	.06	−.11	.07
Canada	−.08	−.01	−.24	−.06	−.39
Denmark	−.22	.01	.01	−.05	−.25
Finland	.06	−.08	−.15	.11	−.06
France	.13	.06	.04	−.01	.22
Germany	.12	.05	.11	−.01	.27
Greece	.06	.01	−.09	−.24	−.26
Ireland	.15	.23	.09	.07	.54
Italy	.25	.13	−.13	−.17	.08
Japan	.13	−.11	−.08	.00	−.06
Netherlands	−.14	.06	.06	.02	−.01
Norway	−.11	−.01	−.01	.13	.00
Spain	−.10	.15	.07	−.35	−.23
Sweden	−.08	−.15	−.07	−.02	−.32
Switzerland	.03	.03	−.01	.06	.10
UK	.01	.14	.02	−.06	.11
US	−.05	.00	−.13	−.10	−.27

We think that a reasonable criterion for a "puzzling episode" is an entry that exceeds .15 in absolute value. This criterion leads to 11 such episodes.[7] Of these, four of them involve hours decreasing by too much, while the remaining seven cases hours do not decrease enough. Although we do not present the detailed information, we note that if one carries out this same exercise in the benchmark model but allowing for government consumption to be valued by consumers, then there are 23 such episodes, and each one is characterized by hours decreasing too much. It follows that the inclusion of distorting taxes in the analysis

[7] One of these is Denmark in the period 1960-1970, which may be due to the data issue noted earlier in the paper.

not only greatly reduces the set of episodes considered puzzling, but also very much changes the nature of the puzzling episodes.

If one applies the same criterion to the aggregate changes over the period, then there are ten countries with changes that exceed .15 in absolute value. Of these, seven are cases in which the issue is that hours did not drop as much theory predicts. It is rather striking to note how the inclusion of taxes alters the important question to be resolved. In our benchmark analysis that abstracted from taxes, we found that all episodes of substantial decreases in hours were puzzling from the perspective of the theory. After including taxes, the puzzling episodes are dominated by cases where hours worked are not decreasing enough.

6.3 Reduced Form Analysis

There are many institutional, policy, and regulatory factors other than taxes that are typically thought to be important factors in the determination of hours of work. Incorporating many of them into the analysis requires substantial extensions of the model, and do not produce simple analytical expressions. Incorporating these elements into the analysis is therefor left for future work.

However, we felt that a simple reduced form exercise might be valuable in giving a crude sense of the possible importance of these other factors. The idea of this analysis is to check which factors might be highly correlated with the wedges that are produced from our benchmark analysis. Loosely speaking, any factors that are not highly correlated with the wedge factors are probably not likely to be of great importance in accounting for changes in hours worked over time. Our strategy is to perform panel regressions to investigate the importance of taxes and indices of labor market institutions in explaining the evolution in wedges. We stress that our exercise is not intended to provide estimates of deep parameters, but simply some *prima facie* evidence regarding the potential importance of various factors.

With respect to the institutional variables, we refer the reader to the Labour Market

Institutions Database constructed by Nickell and Nunziata (2001) and collect the following variables:

1. EP = index for Employment Protection
2. UDNET = Net Union Density
3. CO and COW = measures of bargaining coordination
4. BRR = Benefit Replacement Ratio
5. BD = Benefit Duration

These variables have been largely used in the literature for their ability to capture important relations between institutions and labor market outcomes. Our specification[8] is

$$\log(\Delta_{it}) = \alpha_i + \beta \log(\tau_{it}) + \gamma' X_{it} + \varepsilon_{it}$$

where Δ_{it} is the time series of wedges, α_i is a country fixed effect, τ_{it} is the tax rate used in our analysis, and X_{it} includes the institutional regressors. The sample covers the years 1975-1995 and includes all countries considered in our analysis except New Zealand and Greece. The wedge series that we use in these regressions corresponds to the values calculated when we include government consumption in our measure of household consumption. Table 6 reports our estimation results.

[8] We experiment using variables in levels instead of logarithm. Results are not affected.

Table 6

Regression Results: Dependent Variable is log (Δ_{it})

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tax	.73(.05)	.69(.05)	.76(.05)	.71(.05)	.77(.06)	.78(.06)	.60(.06)	.68(.06)
EP		.13(.04)						.10(.04)
UD			-.21(.09)					-.33(.09)
CO				-.42(.02)				.00(.02)
COW					-.03(.01)			-..04(.01)
BRR						.11(.05)		.09(.04)
BD							-.22(.05)	-.26(.05)
R-sq	.36	.38	.37	.37	.37	.37	.40	.46

There are three main findings. First, the taxes are able to account for a significant fraction of the variation in wedges. Second, this finding is unaffected by the inclusion of any of the other factors, either individually or collectively. Third, although many of the other factors are statistically significant, they add relatively little explanatory power either individually or collectively.

While these reduced form regressions can only be interpreted as identifying correlations and partial correlations, we think it is interesting that of the factors considered, taxes appear to be the most highly correlated.

7 Summary and Conclusion

We have used the neoclassical growth model to shed light on the large reduction in hours worked observed in the OECD over the last 50 years, and the large variation in the magnitude of this reduction across countries. The standard growth model without distortions is able to account for virtually none of the observed reduction in hours worked. In contrast, a model that allows for tax distortions as found in the data is able to account for virtually all of the average reduction across countries. Looking at individual countries we see that in some cases the model implies too large of a reduction in hours, while in others it implies too small of a reduction. For example, the large reduction in hours in Ireland in the period

1956-1980 is not accounted for by the theory, nor is the large increase in hours of work in Canada and the US during the last 20 years. Future work should focus on understanding what additional factors might account for these discrepancies between the theory and data.

There are several avenues which we think are likely to be important. First, our analysis may have abstracted from important features of the tax and transfer schemes in place. For example, Ragan (2005) and Rogerson (2005) argue that the effect of tax distortions in Scandinavian countries is partly undone by the fact that tax revenues are used to subsidize market activities such as child and elderly care. Second, our analysis has abstracted from home production. Recent work by Aguiar and Hurst (2005) shows that the increase in market work in the US over the last 40 years reflects a decrease in time spent in home production and not a reduction in leisure. Freeman and Schettkat (2006) show that time devoted to home production in many European countries is higher than in the US. The lack of time series data on home production precludes one carrying out our exercise in a model that explicitly allows for home production, but assessing the impact of home production on our findings is important. Third, it is important to assess the role of other factors in accounting for the wedges that we have measured.

8 Appendix : Data

Data on consumption and GDP are taken from International Monetary Fund's *International Financial Statistics*. Consumption is the series xxx96F.CZF or xxx96.ZF (xxx stands for country code). Government consumption is the series xxx91F.CZF or xxx91.ZF. GDP is the series xxx99B.CZF or xxx99B.ZF. These series are at current (national) prices and refer our benchmark case. We also experiment using real series and, in order to take into account changes in the relative price, we construct real consumption and real GDP by using consumption price deflator and GDP deflator as calculated by OECD. Real variables calculated using IFS series on GDP deflator and CPI produced similar results. Our findings are not affected by the use of real instead of nominal series, but the sample in the former case is shorter (price series are not available for the whole period for all countries).

Data on employment and population are from OECD's *Economic Outlook* and *Main Economic Indicators.* Series for hours worked are from Groningen Growth and Development Centre (GGDC) and The Conference Board. For population we consider working age population (between 15 and 64 years old), while hours is the total number of hours worked over the year divided by the average numbers of people in employment.

In our regressions, we use as measures of institutional variables the indices available from the Nickell-Nunziata Labor Market Institutional database (2001). We invite the reader to consult the original source for the exact details behind the construction of these variables.

References

[1] Bencivenga, Valerie R., (1992),"An Econometric Study of Hours and Output Variation with Preference Shocks", *International Economic Review*, 33(2), pp. 449-71.

[2] Carey, D. and J. Rabesona (2002), "Tax Rates on Capital, Labour and Consumption", *OECD Economic Studies* No 35 2002/02, Paris.

[3] Chari, V.V., Kehoe Patrick J. and McGrattan Ellen R., (2004), "Business Cycle Accounting", *NBER Working Paper* 10351.

[4] Cole, Harold and Ohanian, Lee E. (2002) "The Great U.S. and U.K. Depressions through the Lens of Neoclassical Theory, *American Economic Review.*

[5] Freeman, R., and Schettkat R., "Marketization of Household Production and the EU-US Gap in Work", Economic Policy 20 (2005), 6-50.

[6] Gali', Jordi, Gertler, Mark and Lopez-Salido, J. David, (2002), "Mark-ups, Gaps, and the Welfare Cost of Business Fluctuations", *NBER Working Paper* 8850.

[7] Hall, Robert E, (1997), "Macroeconomic Fluctuations and the Allocation of Time", *Journal of Labor Economics,* January, vol. 15, no. 1, pp. 223-250.

[8] Ingram, Beth, Kocherlakota, Narayana and Savin, N.E., (1994), "Explaining Business Cycle: A Multiple Shock Approach", *Journal of Monetary Economics*, 34 (3), 415-428.

[9] Maddison, A., *Monitoring the World Economy 1820-1992, OECD, Paris, 1995.*

[10] Mendoza, Enrique G., Razin, Assaf and Tesar, Linda, (1994), "Effective Tax Rates in Macroeconomics: Cross-Country Estimates of Tax Rates on Factor Incomes and Consumption", *Journal of Monetary Economics*, 34, pp. 297-323.

[11] Mulligan, Casey B., (2002), "A Century of Labor-Leisure Distortions", *NBER Working Paper 8774.*

[12] Nickell, S. and L. Nunziata (2001), "Labor market Institutions Databse", Center for Economic Performance.

[13] Parkin, Michael, (1988), "A Method for Determining Whether Parameters in Aggregative Models are Structural", *Carnegie-Rochester Conference Series on Public Policy*, 29, pp.215-52.

[14] Prescott, Edward C., (2004), "Why Do Americans Work So Much More than Europeans?", *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 28, No. 1, July, pp. 2-13.

[15] Ragan, K., "Taxes, Transfers and Time Use: Fiscal Policy in a Household Production Model," mimeo, University of Chicago, 2005.

[16] Rogerson, Richard, (2005), "Taxation and Market Work: Is Scandinavia an Outlier?", *Working Paper*, Arizona State University.