Empirical Evidence on Growth and Busienss Cyles

Martin Zagler*

January 31, 2014

Abstract

A new technology is a bold new combination of production factors that potentially yields a higher level of total factor productivity. The optimal combination of input factors is unknown when an innovation is pursued. A larger targeted innovation may require a greater change in the optimal combination of production factors employed and increases volatility alongside with economic growth. This paper argues that the current crisis, like many before, has been caused by innovation (here in the financial sector). After presenting a model, the paper empirically investigates the relationship between long-run economic growth and output volatility for the time series experience of 21 OECD countries between the years 1961 and 2005. After applying a pooled OLS estimator and a series of robustness checks, we conclude that there is strong empirical evidence for a positive relationship between output variability and economic growth.

Keywords: Growth, Volatility, Cycles, Innovation

JEL-Codes: E32, O33, O47

1 Introduction

An innovation is the active pursuit of entrepreneurs out of a profit motive to find a new and more productive technology. This new technology may require a different combination of inputs in production. The uncertain change in the optimal combination of production factors generates a cost for firms, if they are required to write employment contracts one period in advance. If firms set an optimal combination of production factors different from the optimal combination of production factors, output will fall below its potential level. This will result in lower revenues, which can be interpreted as volatility costs. Both economic growth and volatility are thus produced endogenously by firms decisions, and, in contrast to traditional theories of economic growth and the business cycle, can therefore be instrumentalized by economic policy.

The driving force for growth is innovation and technical progress. In this respect, it does not differ from existing theories of economic growth. (for a

*Corresponding author: Martin Zagler, Universita del Piemonte Orientale and WU Vienna University of Economics, martin.zagler@gmail.com

survey, see Aghion and Howitt 1998). In contrast to existing theories of economic growth, this paper differs in identifying a different boundary to economic growth. Previous models of economic growth have focused on accumulation (Solow 1956, Harrod 1948, Domar 1946, Rebelo 1991, Romer 1986, Lucas 1988, Barro 1990) and on resource constraints (Romer 1990, Grossman and Helpman 1991, Aghion and Howitt 1992). In the prior, growth was bound as (capital) accumulation could only be finite. In the latter, growth was bound by the amount of workers available in the innovation process. Indeed, the dependence of growth on resources led to the Jones critique (1995), which essentially states that current growth rates are unsustainable, as they have required an ever increasing number of workers in research and development (R & D). Here, by contrast, growth will be constrained by the amount of risk entrepreneurs are willing to accept. As the cost associated with innovation risks increases exponentially, there appears a natural boundary to economic growth that is not related to the amount of resources devoted to the growth process.

Despite the novelty of the approach, the paper is related to several strands of literature. Technological change as a source of economic growth has been addressed early on in growth theory. The rate of technical progress is the only source of long-run growth in per capita GDP in the Solow model (1956). Still, technical progress only affected total factor productivity (TFP), but held the optimal factor input combination for given prices constant. Harrod (1953) suggested that technical progress is embodied entirely in the factor labor as the only possibility consistent with the facts. Thus the "great ratio" of capital to efficiency units of labor remains constant along the balanced growth path. However, labour augmenting technical progress leads to a permanently growing wage, and hence a constant but foreseeable shift in the optimal factor input combination away from labor and toward capital.

The first wave of endogenous growth models, which all have in common constant returns to scale with respect to reproducible factors of production, retain the "great ratio" properties of factor shares (Rebelo 1991, Romer 1986, Lucas 1988). In principle, one could introduce a cost of adjustment to a higher level of production. If these costs are exponential and significant, the economy may grow at a permanently lower growth path, or the growth rate may increase or decrease over time, in accordance with the specific properties of adjustment costs. Under a balanced growth path, the factor shares will remain constant even under these assumptions. In some respect, these adjustment costs mimic the costs of employing a suboptimal factor mix, as proposed here. However, no model of adjustment cost would be able to generate a cyclical behavior of the economy.

Technical change that favors one factor of production over another was later introduced under the heading of skill-biased technical change (Ter Weel and Sanders 2000), in order to explain the growing wage gap between skilled and unskilled workers (Katz and Murphy 1992). Of the several hypotheses brought forward in the debate, the one closest related to this paper simply claimed that technical change somehow favored one factor over another (Acemoglu 1998). This paper, too, allows for change in relative wages between two groups of workers. However, here the movement of relative wages need not be unidirectional and certainly will not be predictable.

None of the above papers have been able to draw the attention to the common determinants of economic growth and economic cycles. However, there is a number of empirical papers suggesting a relationship between economic growth and cycles. Campbell and Mankiw (1987) were amongst the first to report permanent effects on the level of GDP from shocks to output growth, first for the US and later on for a selected sample of various countries (Campbell and Mankiw 1989). Hall (1988) and Burnside, Eichenbaum and Rebelo (1993) show that the Solow residual is correlated economic variables, and can therefore not be purely exogenous, as suggested by the real business cycle literature, suggesting that trend and fluctuation of output should be investigated jointly.

Several authors have attempted to model the joint determination of growth and cycles. A noteworthy first attempt was by King, Plosser, and Rebelo (1988). They argue that the business cycle indeed affects augmented factor productivity through the quality of capital and labor, capacity, energy and natural resources, foreign trade and structural effects. Nested within real business cycle theory, temporary stochastic shocks only cause temporary deviations from potential output. Given on average positive productivity gains, trend growth itself is stochastic but remains exogenous. While the fluctuations around average movements in productivity can be interpreted as the business cycle, movements in trend itself remain unexplained within the theory of real business cycles.

As business cycle shocks are exogenous, policy can only make things worth. In the model presented below, where both growth and volatility are determined endogenously by the choice of economic actors, policy will be influential. We will illustrate this with an important special case. Using a model of a small open economy, we can show source based capital income taxation can alter the trade-off between economic growth and volatility. An increase in capital income taxation will reduce the economies proneness to volatility and stabilize the economy, albeit on a slower growth path. Capital income taxation therefore plays a role in this model, inducing stability. This is in stark contrast to conventional models of capital taxation, where one typically finds that with fully mobile capital, taxing capital will be inferior to taxing immobile labor (Sinn, 2003). This paper thus suggest a role for capital income taxation, the stability motive, and prescribes it the role of an automatic stabilizer.

There are several recent papers that try to model both growth and cycles endogenously. Matsuyama (1999) and Waelde (1999) argue that changes in productivity happen only sporadically, either because there is a stochastic element of failure intrinsic in innovation, or because firms prefer to invest in capital accumulation after periods of high productivity growth. Within an elaborate endogenous growth model, Aghion et al. (2005) investigate the effect of exogenous shocks on growth and volatility. The interesting feature is that policy choices (in their case concerning credit market regulation) may work on the trade-off between economic growth and fluctuations. Comin and Mulani (2006) present an innovation model, with firm specific and general innovations. The prior lead to volatility, the latter to economic growth. Here, too, the growth and volatility are endogenous, and the trade-off depends primarily on market structure. Closest to this proposal is a recent paper by Bojanovich (2005). There the choice of a growth rate leads to a positively correlated stochastic cost. However, Bojanovich fails to motivate the source of the stochastic cost.

For a long time, the field of macroeconomics has between firmly divided between the analysis of the business cycle and the investigation of long-run determinants of economic growth. This distinction, however, is rather arbitrary and has been challenged by recent theoretical models and by empirical evidence that points to long-run performance being explained in part by business-cycle behavior and output variability. The aim of this paper is to empirically investigate the relationship between economic growth and output volatility.

The earliest theoretical argument for a relation between economic growth and the business cycle dates back to Schumpeter (1939), who argued that recessions provide a cleansing mechanism for the economy, where old technologies get replaced by newer technologies, and will be better adapt to economic growth thereafter. In a similar spirit Black (1981) argues that the average severity of a society's business cycle is largely a matter of choice. His idea was that economies face a positive risk-return trade-off in their choice of technology, as economic agents would choose to invest in riskier technologies only if the latter were expected to yield a higher return and hence, greater economic growth.

A series of papers have subsequently focused on the relationship between volatility and growth in exogenous growth models. On the one hand, the focus was on the impact of volatility on uncertainty, precautionary savings and hence accumulation of capital (cf. Boulding (1966), Leland (1968), Sandmo (1970)). On the other hand, Bernanke (1983) and Pindyck (1991) argue that if there are irreversibilities in investment, then increased volatility will lead to lower investment and hence lower capital accumulation. Both strands of literature have in common that they are based on exogenous growth models, hence whilst there may be transitional changes in growth rates due to changes in volatility, in the long-run economic growth will be exogenous.

More recently, in an endogenous growth model Aghion & Saint-Paul (1993) and Aghion et al. (2005) show that the sign of the relation depends on whether the activity that generates growth in productivity is a complement or a substitute to production. In the case where they are substitutes, since the opportunity cost of productivity-improving activities such as reorganizations or training falls in recessions, larger variability leads to higher long-term growth. This idea has recently been formalized in an endogenous growth framework by Jovanovich (2006).

A number of empirical studies on the relationship between growth and volatility has been conducted. Campbell & Mankiw (1987) were amongst the first to report permanent effects on the level of GDP from shocks to output growth, first for the US and later on for a selected sample of various countries (Campbell & Mankiw (1989)). Whilst it provides a confirmative test for models of exogenous growth and volatility, these studies fail to provide a test for models of endogenous growth and volatility.

The first empirical study that can be applied to endogenous growth models was done by Zarnowitz (1981). He identified periods of relatively high and relatively low economic stability by reviewing annual real GDP growth rates in the U.S. between 1882 to 1980 and accounts found in the literature on economic trends and fluctuations. He then calculated the yearly growth rate and the variance of the periods with high economic stability (group A) and low economic stability (group B). Though the mean growth rate of group A was higher, he could not reject the null hypothesis that the difference between the mean growth rates for groups A and B was due to chance.

The first econometric study investigating the link between growth, output variability—as measured by the standard deviation of the growth rate—and further macroeconomic variables was conducted by Kormendi & Mequire (1985). By averaging each country's time series experience into a single data point and estimating a cross-section of forty-seven observations, they found that higher output variability leads to higher economic growth. Grier & Tullock (1989), who used a pooled structure (five-year averaging) to account for both between- and within-country effects, confirmed Kormendi and Meguire's results.

The paper closest to ours is by Mills (2000). He applied various filters that are explicitly designed to capture movements in a time series that correspond to business-cycle fluctuations in twenty-two countries. Subsequently, he calculated the standard deviation of the output (filtered) series and visualized the bivariate relationship between growth and volatility by superimposing robust nonparametric curves on scatter plots. He found a positive relationship. In contrast to our paper, Mills (2000) suppresses all fluctuations of output at frequencies higher than his filter.

When analyzing the relationship between economic growth and output fluctuations, we are essentially investigating the first moment of the time series in first differences, and its corresponding second moment over the mean, i.e. the variance of the differentiated time series. There exists a standard econometric tool to analyze this relationship, the generalized auto-regressive conditional heteroscedacity (GARCH) class of models. And indeed, several authors have employed this methodology to analyze the relationship of output and volatility.

Ramey & Ramey (1995), using a panel structure, measured volatility as the standard deviation of the residuals in a growth regression consisting of the set of variables identified by Levine & Renelt (1992) as the important control variables for cross-country growth regressions. Ramey & Ramey (1995) use the estimated variance of the residuals in their regression, under the assumption that it differs across countries, but not time. In such, it can be considered an early predecessor of GARCH models[1]. They find a negative relation between long-run growth and volatility. By contrast, Caporale & McKiernan (1998) and Grier & Perry (2000) examined the issue from a pure time series perspective. Caporale & McKiernan (1998) ran an ARMA(1,2)-GARCH(0,1)-M model and Grier & Perry (2000) ran a complex bivariate GARCH(1,1)-M model for U.S. GDP growth. The former found a significant positive relationship while the latter found an insignificant positive relationship between growth and volatility.

The fact that these studies yield opposite results may come as a surprise. However, GARCH models were invented for financial time series, with a large number of observation. In Monte-Carlo simulations, presented in appendix A, we demonstrate that the widely-used and highly-sophisticated GARCH-in-mean models are inappropriate for this purpose as they require the estimation of too many parameters for the short time series that normally confront economists.

This leaves us with the more conventional approach of separating the time series into a trend and a cyclical component, and then investigate their relationship. There is a large number of filters available, most of them developed by the finance literature. We have decided to adopt the HP-filter. Our measure of volatility is superior to any other measure of volatility we investigated due to its stability with respect to small changes in the data.

The empirical analysis presented here is based on the growth experience of twenty-one OECD countries between 1961 and 2005. After calculating the trend growth rate for each country using the HP-filter, we divided the data for each country into three, fifteen-year, non-overlapping sub-samples. For each

[1] With a single estimate per country, we cannot simulate their results as done in A

sub-sample, the average growth rate and the volatility–based on the squared deviations of the actual growth rate from the trend growth rate–was computed. This not only mitigated the effect of assuming constant volatility and constant growth rates, the technique also accounted for the within-country variation of the volatility in our subsequent regression analysis. After running a series of robustness tests, we conclude that there is a significant positive relationship between output variability and growth. This relationship is robust against outliers and does not hinge on the sub-sample period chosen.

2 The Model

We will analyze the relationship between economic growth and volatility in a partial equilibrium[2] model of a small open economy[3], where capital is fully mobile internationally with a world market prize of ρ, whereas labor is fully immobile and comes at fixed supply $L = 1$. Aggregate output is assembled by homogeneous inputs from n firms[4], with productivity of the assembly equal to A_t,

$$Y_t = A_t \sum_{i=1}^{n} y_{i,t} \tag{1}$$

Each firm in the economy strives to gain a competitive advantage over others by implementing new technologies and rendering the factor labor more efficient. These will yield a labor augmenting productivity gain of $a_{i,t} > 1$, which will last for one period[5]. There are no direct costs associated with this productivity gain, however, firms will face uncertainty over the optimal factor input combination, which is increasing in the size of the productivity gain[6]. These individual productivity gains will generate non-appropriable public knowledge[7] that is used in the assembly of the output good (1), so that the average of all $a_{i,t}$ will be the rate of technical progress in the economy,

$$A_t = \frac{A_{t-1}}{n} \sum_{i=1}^{n} a_{i,t} \tag{2}$$

An new technology will only have a transitory effect for the single firm, but will have a permanent effect on the economy on the whole. We can think of any successful innovations, or implementations of new technologies, as being copied with a one period lag by all other firms in the economy. Firm specific technological knowledge thus turns into general knowledge, with the added advantage that the uncertainty about the optimal factor input combination will vanish,

[2]It should not be very difficult to extend the model to a full equilibrium model.

[3]Wildasin (1995) shows that this assumption is not fully innocent, as a shift from a closed to an open economy may shift risk from capital to labor.

[4]The number of firms n is exogenously given, but could well be determined by product market regulations and national competition policy. We will analyze changes to the number of firms in this light below.

[5]We can think of $a_{i,t}$ as efficiency gains. If the firm engages in productivity enhancing activities, they will extract labor efficiency above unity, otherwise not.

[6]As shown below, this will induce firms to choose a finite level of productivity increases

[7]This is a knowledge externality typical for endogenous growth models.

too. This is in stark contrast to the innovation literature, where patent protection for innovation lasts forever (Grossman and Helpman 1991) or at least until a new innovation comes around (Aghion and Howitt 1992), but coincides with actual patenting practice and open source innovations. Clearly it applies more to process innovations than to product innovations[8].

Aggregate economic growth will therefore be driven by two sources, disembodied technical progress (2) and output growth of individual firms (3). As the prior will have no impact on volatility, all cyclical components will derive from the later term. This is in accordance with Comin and Mulani (2006), who postulate that firm specific knowledge predominantly drives volatility, whereas general knowledge is responsible for economic growth. Firms[9] produce output with a constant elasticity of substitution technology,

$$y_{i,t} = (1-\phi_{i,t})k_{i,t}^{\sigma} + \phi_{i,t}(a_{i,t}l_{i,t})^{\sigma} \tag{3}$$

where $k_{i,t}$ and $l_{i,t}$ are capital and labor, respectively. We assume that firms must hire capital and labor at the beginning of the period, and before any shock realizes[10]. The elasticity of substitution is given by $1/(1-\sigma)$. In order to ensure substitutability between production factors, we must have $\sigma < 1$. In this case, the production function exhibits decreasing returns to scale, with a scale factor equal to σ. In order to ensure a positive marginal product for both capital and labor, this requires $0 < \sigma < 1$.

We will introduce the above mentioned uncertainty over the optimal factor input combination by assuming that the parameter determining factor shares, $\phi_{i,t}$, changes with the size of the innovation. To simplify matters, we assume that $\phi_{i,t}$ has a bivariate distribution that depends on the size of the technological innovation implemented by the firm,

$$\phi_{i,t} = \begin{cases} \phi_{i,t-1} + \frac{a_{i,t}-1}{a_{i,t}}(1-\phi_{i,t-1}) & \text{with probability } p; \\ \phi_{i,t-1} - \frac{a_{i,t}-1}{a_{i,t}}\phi_{i,t-1} & \text{with probability } 1-p. \end{cases} \tag{4}$$

Several things are worth mentioning at this point. With probability p, firms are hit by a "positive" shock, i.e. the factor share parameter $\phi_{i,t}$ will be larger than before, and with probability $1-p$, firms are hit by a "negative" shock, i.e. the factor share parameter $\phi_{i,t}$ will be larger than before. We shall assume that whether we are faced with a positive or a negative shock to ϕ is drawn once for the entire economy, in order to exclude pooling of resources of groups of large firms. Note that if firms choose no innovation, firm productivity will be $a_{i,t} = 1$, and $\phi_{i,t}$ will remain at the previous level $\phi_{i,t-1}$ irrespective of the realization of the shock. By contrast, if firms choose an infinite level of innovation, $\phi_{i,t}$ will go to unity with probability p and to zero with probability $1-p$. In that case, either the entire amount of capital or the entire amount of labor employed will be completely unproductive and therefore only costly for the firm. At this point, it may be interesting that the one period ahead expected value of $\phi_{i,t}$ takes the form

[8] For this reason, we have refrained from modeling differentiated products in the first place.

[9] We assume that the number of firms n is large, so that firms need not consider the choice of others in their optimization problem.

[10] We will normalize the price of input goods to unity. This implies that the price of the output good (1) will equal $1/A_t$ and falls with technical progress.

$$E_{t-1}[\phi_{i,t}] = [\phi_{i,t-1} + p(a_{i,t} - 1)]/a_{i,t} \tag{5}$$

If $\phi_{i,t-1} = p$, the expectation of $\phi_{i,t}$ is equal to its previous value. If $\phi_{i,t-1} > p$, the expected value of $\phi_{i,t}$ will be below its previous value and vice versa. We thus have mean reversion in the shock, and the long-run expectation will be $\phi_{i,t-1} = p$.

Firms will choose capital, labor, and the level of productivity in order to maximize expected profits, taking market prizes for their product, capital and wages as given. Firms will pay a source based tax on capital income equal to τ. The first order condition with respect to labor states that the expected marginal product of labor will be equal to the domestic wage,

$$\sigma a_{i,t}^{\sigma} E_{t-1}[\phi_{i,t}] l_{i,t}^{\sigma-1} = w_t \tag{6}$$

In equilibrium, labor demand of all n firms must equal supply, which has been normalized to unity. Labor demand is downward sloping, so the wage can ensure equilibrium in the labor market. The firm's decision to innovate will influence wages. Substituting the expected value of the factor share parameter $\phi_{i,t}$ from equation (5), we note that wages will increase with the size of the productivity gain $a_{i,t}$. The first order condition with respect to capital reads,

$$\sigma E_{t-1}[1 - \phi_{i,t}] k_{i,t}^{\sigma-1} = (1 + \tau)\rho \tag{7}$$

and holds that the marginal product of capital should equal the gross interest rate $(1 + \tau)\rho$. Ceteris paribus, an increase in the tax on capital (or the world interest rate) will reduce the domestic demand for capital, given $\sigma < 1$. So capital taxation will lead to capital flight, and it has been proven elsewhere (Sinn 2003) that this has negative consequences both for the level and the distribution of national income. As the change in the price of the domestic capital stock will alter the optimal factor input combination, firms will wish to alter factor share parameter $\phi_{i,t}$ by changing the rate of innovation. And it is this link which will ensure the role of capital income taxation on the trade-off between growth and volatility in this economy. The optimal degree of technical change is given by

$$\frac{E_{t-1}[\phi_{i,t}]}{a_{i,t}} \left[\sigma(a_{i,t} l_{i,t})^{\sigma} + \epsilon(a_{i,t} l_{i,t})^{\sigma} - \epsilon k_{i,t}^{\sigma}\right] = 0 \tag{8}$$

where the elasticity of the factor share parameter with respect to changes in productivity ϵ is defined as

$$\epsilon = \frac{\partial E_{t-1}[\phi_{i,t}]}{\partial a_{i,t}} \frac{a_{i,t}}{E_{t-1}[\phi_{i,t}]} = \frac{p - \phi_{i,t-1}}{\phi_{i,t-1} + p(a_{i,t} - 1)}$$

There are three distinguished effects in the first order condition with respect to technical progress (8). The first is the direct effect of technical progress on labor productivity. The second and the third are indirect effects of technical progress on labor and capital productivity, and it is always positive. The two indirect effects have opposite signs. With capital productivity higher than on average, $1 - \phi_{i,t-1} > 1 - p$, the elasticity of the factor share parameter with respect to changes in productivity ϵ is positive. Thus one can expect a decrease in capital productivity and an increase in labor productivity, and vice-versa[11].

Substituting the first (6) and second (7) first order conditions into the third (6), we obtain a solution that depends on a single firm specific variable, the rate of technical progress $a_{i,t}$. Given $a_{i,t} \leq 1$, and under fairly mild parameter restrictions, we find that there is a unique solution for technical progress $a_{i,t} = a_t$, and all firms will make the same choice of innovation. This implies that they will all choose the same capital stock k_t and the same number of employees, $L_{i,t} = L_t$, equal to $1/n$ due to labor market clearing. Substituting labor market clearing and the second first order condition (7) into the third (6), we obtain an implicit solution for the degree of technical progress[12],

$$\sigma n^{1-\sigma} a_t^{\sigma-2}[a_t(1-p) + p - \phi_{t-1}] \left(\frac{\epsilon}{\sigma+\epsilon}\right)^{\frac{1-\sigma}{\sigma}} = (1+\tau)\rho \tag{9}$$

This gives a unique and finite solution for technical progress a_t. Economic growth is bounded not by capital accumulation or innovation costs, which have been assumed to be zero. The capital stock will not grow without bound in this economy, and is bounded from above if $\phi_t = 1$ following equation (7). The reason for bounded economic growth in this model is the fact that infinite growth would yield infinite costs due to a mismatch in the optimal factor input combination. Firms will therefore prefer to induce finite technical change to avoid exuberant costs.

3 Growth and Cycles

With a positive rate of technical progress a_t, the economy will exhibit volatility. Individual firms contribute to aggregate output (1) through its production y_t and through its contribution to public knowledge a_t. We will therefore use the measure $a_t y_t$ as a measure for the economy as a whole. In the case of a positive shock to the factor share parameter ϕ_t, the aggregate output share of a particular firm will equal

$$a_t y_t^+ = (1-\phi_{t-1})k_t^\sigma + \phi_{t-1}(a_t l_t)^\sigma + (a_t - 1)(a_t l_t)^\sigma \tag{10}$$

By contrast, if the factor share parameter is hit by a negative shock,

$$a_t y_t^- = (1-\phi_{t-1})k_t^\sigma + \phi_{t-1}(a_t l_t)^\sigma + (a_t - 1)k_t^\sigma \tag{11}$$

The difference between those two states is a good measure of volatility in the economy, and is equal to

$$a_t(y_t^+ - y_t^-) = (a_t - 1)[(a_t l_t)^\sigma - k_t^\sigma] \tag{12}$$

It is important to note that volatility is monotonically increasing in the rate of technical progress a_t. Expected output is equal to

[11]Note that with $\phi_{i,t-1} = p$, we have $\epsilon = 0$ and the two indirect effects exactly offset each other. This would induce firms to set an extremely high rate of innovation, only to subsequently return toward the long-run value of the factor share parameter. As the probability of $\phi_{i,t-1} = p$ is zero, we shall rule out this case by assumption, e.g. set $\phi_{i,t} = 1/2$ if and only if $\phi_{i,t-1} = p$, with $p \neq 1/2$.

[12]The solution is unique, as argued above.

$$E_{t-1}(y_t) = \frac{1}{a_t}[1 - \phi_t + (a_t - 1)(1 - p)]k_t^\sigma + \frac{1}{a_t}[\phi_t + (a_t - 1)p](a_t l_t)^\sigma \quad (13)$$

which is a weighted average between output under the current factor share parameter ϕ_t and the expected long-run distribution parameter p, where a larger rate of technical progress a_t puts less weight on the current value. Costs in this economy are given by labor and capital costs

$$w_t L_t + (1 + \tau)\rho k_t = \sigma E_{t-1}(y_t) \quad (14)$$

This allows us to determine profits in both states of the world,

$$\pi_t^+ = \frac{1-\sigma}{a_t}[(1-\phi_{t-1})k_t^\sigma + \phi_{t-1}(a_t l_t)^\sigma] + \frac{a_t - 1}{a_t}[(1-\sigma p)(a_t l_t)^\sigma - (1-p)\sigma k_t^\sigma] \quad (15)$$

which is a weighted average between profits in the absence of innovation and profits due to innovation, where the latter can be negative. Similarly, profits in the other state of the world equal

$$\pi_t^- = \frac{1-\sigma}{a_t}[(1-\phi_{t-1})k_t^\sigma + \phi_{t-1}(a_t l_t)^\sigma] + \frac{a_t - 1}{a_t}[(1-\sigma(1-p))k_t^\sigma - \sigma p(a_t l_t)^\sigma] \quad (16)$$

where the second part can once again be negative. Unless σ is very close to unity, we can ensure positive profits in both states of the world. Expected profits equal

$$E_{t-1}(\pi_t) = (1-\sigma)[[(1-p)k_t^\sigma + p(a_t l_t)^\sigma] + \frac{\phi_{t-1} - p}{a_t}[(a_t l_t)^\sigma - k_t^\sigma] \quad (17)$$

which is bigger than profits under no innovation, and thus ensures a positive rate of innovation in the economy.

4 Government Policy

The implicit solution (9) allows for ample policy analysis. We will look at three distinct policy experiments, an increase in the source based capital income tax τ, automatic stabilizers, and an increase in product market liberalization, modeled through an increase in the number of firms n. We will also look at other forms of taxation in order to see why this model, as opposed to conventional theories, will give a role to source based capital income taxation, even in a world with perfectly mobile capital.

The first policy experiment that we introduce is an increase in the source based capital income tax. Whilst we cannot take a derivative of technical progress a_t with respect to the tax rate, we can do the opposite,

$$\frac{\partial(1+\tau)}{\partial a_t} = -\frac{1}{\rho} n^{1-\sigma} a_t^{-\frac{2}{\sigma}} \left(\frac{p - \phi_t}{p}\right)^{\frac{1-\sigma}{\sigma}} [a_t(2-2\sigma)(1-p) + (2-\sigma)(p-\phi_t)] \quad (18)$$

For values of $\phi_t < p$, the solution implies that in increase in capital income taxes will reduce the innovative effort and firms, and thus reduce economic growth and volatility. In the special case of $\phi_t = p$, the solution to the above equation 18 is zero, and we cannot invert the result for the impact of taxation on technical progress. For values of $\phi_t > p$, things turn tricky. We can run into irrational solutions, which we rule out by assumption. In the case $(1 - \sigma)/\sigma$ is a multiple of 2, we will again find that an increase in capital taxes will reduce technical innovation. Otherwise, with the term in parenthesis negative, the sign of the effect will depend on the sign of the term in square brackets. Clearly, for values of ϕ_t close to p, the term will be positive, and capital income taxation will lead to an increase in innovative activity. When $\phi_t - p > (1 - \sigma)[2a_t(1 - p) + p]$, the relationship inverts again and we obtain a negative relation between capital income taxation and economic growth. A sufficient condition to ensure that the effect inverts for $\phi_t < 1$ is $a_t < \frac{2-\sigma}{2-2\sigma}$. The fraction converges to infinity as the scale parameter σ approaches unity, and will thus be easily satisfied.

A capital income tax will therefore lead to lower growth and higher stability whenever labor productivity is low or very high, and only with medium high labor productivity will it lead to more growth and instability. With the economy equally likely to fall above or below p, policymakers may wish to introduce capital income taxation in order to reduce innovative activities and increase stability.

Noting from equation 9 above, the effect of an increase in the number of firms n is exactly opposite to an increase in capital income taxation. Countries may therefore wish to introduce product market regulations for the very same reason they introduce capital income taxes, in order to increase economic stability.

The automatic stabilizing effect obtained from capital income taxes cannot be easily achieved through other forms of taxation. Labor taxes will only alter the current wage rate, as labor is in constant supply, and change nothing in the relationship between growth and volatility. Residence based capital income taxes would tax worldwide capital income of domestic residents, and insofar as the economy is small, this would not influence world interest rates and innovation. Income taxes, which in this economy would be a combination of labor taxes, source and residence based capital income taxes, would only influence the trade-off through its effect on source based capital income taxes, but require a far larger amount of tax revenues to achieve the same effect. Finally, consumption taxes would tax output consumed at home. Given undifferentiated products, this would not influence domestic suppliers of products to the world market, and therefore have no influence, either.

5 Conclusions

This paper has established a link between economic growth and economic volatility. The idea was that a new technology is a new combination of production factors that yields a higher level of total factor productivity. However, the optimal combination of input factors is unknown when an innovation is pursued. A larger targeted innovation requires a greater change in the optimal combination of production factors employed and increases volatility alongside with economic growth.

Economic growth is bounded by the costs associated with uncertainty over the optimal factor combination. The further the economy will depart from the anticipated optimal factor input combination, the higher will be these factor costs. As these costs are increasing exponentially due to the convexity of the production function, firms will pursue only finite changes in productivity, thus inducing bounded economic growth.

We show that economic policy can interfere in this relationship with by adjusting source based capital income taxes. An increase in capital income taxes will induce a slower targeted level of technical progress, but also lead to lower volatility. Capital income taxes can therefore be used to stabilize the economy, giving a motive why small open economy may still wish to introduce them, despite their negative allocative and distributive effects. No other form of taxation can achieve this goal equally.

6 The Data

The data for this study came from the AMECO database.[13] It is the annual macro-economic database of the European Commission's Directorate General for Economic and Financial Affairs (DG ECFIN). All 21 countries (Australia, Austria, Belgium, Canada, United Kingdom, Finland, France, Greece, Iceland, Ireland, Italy, Japan, Luxembourg, Mexico, Netherlands, Portugal, Spain, Sweden, Switzerland, Turkey, and USA) for which continuous annual series for gross domestic product at constant market prices per capita were recorded for the period of 1960-2005 were used for analysis.

7 Methodology and Results

7.1 Modeling Trend and Volatility

We will investigate time series properties of a particular nature. In order to analyze the relationship between economic growth and volatility, we will ask whether a measure of volatility is correlated with changes in output growth. Several measures for both output growth and the volatility are feasible, and we will discuss them below. Whilst for economic growth, the change in the level of output–maybe averaged over several periods, which would be a trend– is a natural candidate, measures for the business cycle are volatility measures. Volatility refers to the spread or dispersion of all likely outcomes of a random variable. It is often measured as the sample standard deviation. Formally, we investigate a relationship such as,

$$g_t = \kappa + \gamma \sigma_t + u_t \tag{19}$$

where κ is a constant, γ is a parameter, and σ_t measures the standard deviation of the time series[14]. u_t is an error term. For a given time series, one could estimate the above equation (19), then use the estimator for the variance

[13] http://ec.europa.eu/economy_finance/indicators/annual_macro_economic_database/ameco_en.htm

[14] We refrain from including control variables in our estimation. Unless control variables would be correlated with the variance measure adopted, the estimator for γ remains unbiased. Most control variables that we can think of, such as policy variables, would work in favor, reducing the explanatory power of volatility on economic growth.

σ^2 and reestimate the above equation until it converges.[15] This essentially what GARCH models do. Estimating a time-varying variance requires a long time series, a luxury we cannot afford for macroeconomic time series such as GDP. In appendix A, using Monte-Carlo simulations, we show that under reasonable parameter configurations, the variance of the estimator from its true variance is unacceptably large[16].

This leads us to the next best solution of estimating mean and variance separately.[17] The exercises is further complicated as both the mean and the standard deviation are not necessarily constant over time.[18] We will test for constancy over time using three types of unit root tests.

7.2 Unit Root Tests

One clear indication that the assumption of a constant mean and a constant variance of a time series cannot be maintained is when unit root tests point to the non-stationarity of the data. In this case, cross-country regressions based on sample mean and sample variance would lead to bogus results.

Testing for unit roots in the growth rate of GDP using the standard Augmented Dickey-Fuller[19] (ADF) test—with a constant and a trend in the regression equation—results in the failure to reject the null hypothesis of non-stationarity two-thirds of the time (5 % level of significance). Since the way in which classical hypothesis testing is carried out ensures that the null hypothesis is accepted unless there is overwhelming evidence against it and we want to point out that our series are non-stationary, the appropriate way to proceed is to use a test that has the null hypothesis of stationarity and the alternative of a unit root. A test with stationarity as null is the KPSS test. Kwiatkowski et al. (1992) start with the model

$$\begin{aligned} y_t &= \xi t + r_t + \epsilon_t \\ r_t &= r_{t-1} + u_t \end{aligned} \tag{20}$$

where $u_t \sim \text{iid}\left(0, \sigma_u^2\right)$, ϵ_t and u_t are independent, and the initial value r_0 is fixed. The ϵ_t satisfy the linear process conditions of Phillips & Solo (1989) (theorems 3.3,3.14) which allow for all ARMA processes, with either homogeneous or heterogeneous innovations.

The test for stationarity in this model is simply

$$H_0 : \sigma_u^2 = 0 \quad \text{vs.} \quad H_A : \sigma_u^2 > 0 \tag{21}$$

[15]It should also be noted that whenever one has an unbiased estimator for σ^2, the square root of $\hat{\sigma}^2$ is a biased–depending on the shape of the distribution and the sample size–estimator of σ due to Jensen's inequality, $\mathbb{E}\left[\hat{\sigma}\right] = \mathbb{E}\left[\sqrt{\hat{\sigma}^2}\right] < \sqrt{\mathbb{E}\left[\hat{\sigma}^2\right]} = \sqrt{\sigma^2} = \sigma$.

[16]This may be the reason why papers based on this methodology yield contrasting results.

[17]A measure for the spread of a distribution does not necessarily contain all information about its shape, so we can still miss some important features, unless the first two moments (mean and variance) are sufficient statistics to describe the entire distribution.

[18]The analysis by Kormendi & Mequire (1985) basically relies on this assumption.

[19]The number of lags used in the regression is $\text{trunc}\left(\left(\text{length}\left(\text{series}\right) - 1\right)^{\frac{1}{3}}\right) = 3$. This corresponds to the suggested upper bound on the rate at which the number of lags should be made to grow with the sample size for the general ARMA(p,q) setup.

Country	$\mathbf{KPSS}_\mu$	$\mathbf{KPSS}_\tau$	$\mathbf{ADF}_\tau$
Australia	□	□	■
Austria	■	□	■
Belgium	■	□	■
Canada	□	□	■
England	□	□	□
Finland	□	□	■
France	■	■	■
Greece	■	■	■
Iceland	□	□	□
Ireland	□	□	■
Italy	■	□	□
Japan	■	■	■
Luxembourg	□	□	■
Mexico	□	□	■
Netherlands	■	□	■
Portugal	■	□	□
Spain	■	■	■
Sweden	□	■	■
Switzerland	□	□	□
Turkey	□	□	□
USA	□	□	□

Table 1: Unit Root Tests

We performed two tests[20], denoted by KPSS_μ and KPSS_τ based on a regression on a constant μ, and on a constant and a time trend τ, respectively. Even though both tests are very conservative, we reject the stationarity hypothesis in 45% and in 25% of the cases, respectively.

Table 1 shows the results for the ADF test and the two KPSS tests for each country. Black squares denote evidence for non-stationarity (ADF: non-rejection of the null hypothesis, KPSS: rejection of the null hypothesis) while white squares denote evidence for stationarity. Out of our sample of 21 countries, all three tests point to stationarity of the data for only five countries. To summarize, we obtain a dispersed picture, and have to reject the assumption that all series exhibit constancy over time in all countries. We will therefore resort to band-pass filters to identify the trend (growth) and cyclical component of the time series.

7.3 Separating Trend and Volatility

It is often assumed that the time series under investigation, Y_t, can be represented as a weighted sum of periodic functions of the form $\cos(\omega t)$ and $\sin(\omega t)$ where ω denotes a particular frequency:

$$Y_t = \mu + \int_0^\pi \alpha(\omega)\cos(\omega t)\,d\omega + \int_0^\pi \delta(\omega)\sin(\omega t)\,d\omega \tag{22}$$

An *ideal* band-pass filter is a linear transformation of Y_t that isolates the

[20]To estimate σ_u^2 the Newey-West estimator was used.

components that lie within a particular band of frequencies, i.e. the filter only passes frequencies in the range $\omega_L \leq \omega \leq \omega_H$. Applied to GDP growth rates, the filter eliminates very slow-moving ('trend') components and very high-frequency ('noise') components, while capturing intermediate components that correspond to business-cycle fluctuations. The variance of the filtered series, $\hat{g}_t$, could then serve as a measure of volatility.

However, since such an ideal band-pass filter is a moving average of infinite order and therefore requires infinite data, an approximation is necessary for practical applications. Mills (2000) employed the one suggested by Baxter & King (1995) and removed components with frequencies below two years and above eight years.

Building on the graduation method developed by Whittaker (1923) and Henderson (1924), Leser (1961) proposed a filter that is similar to the band-pass, one that has also been widely used in business-cycle research. In economics it is known as the Hodrick-Prescott (henceforth HP) filter. The HP filter is an approximate low-pass filter, i.e. it passes low frequencies but attenuates (or reduces) frequencies higher than the cutoff frequency.

The filtered series is obtained by solving:

$$\min_{\hat{g}_t} \left[\sum_{t=1}^{T} (y_t - \hat{g}_t)^2 + \lambda \sum_{t=2}^{T-1} \left((1-L)^2 \hat{g}_{t+1} \right)^2 \right] \tag{23}$$

where $L^n y_t = y_{t-n} \quad \forall n \in \mathbb{N}$. The first summation term in equation 23 concerns the fit (squared deviations), the second summation term the smoothness (squares of the second differences) of the filtered series. The parameter λ determines the importance of the smoothness relativ to the fit (trade-off). As $\lambda \to \infty$, $\hat{g}_t$ approaches a linear trend.

7.4 Measuring Volatility

We are confronted with the situation whereby some GDP growth series appear to be stationary, while others appear to be trend-stationary, or even non-stationary. In the case of stationarity and trend-stationarity, the growth rate fluctuates around a constant and a linear trend, respectively. In the case of non-stationarity, the growth rate either fluctuates around a deterministic non-linear trend or a stochastic trend. Using different procedures to calculate the variance for each country could inadvertently result in data mining; therefore, we uniformly applied the same variance-extracting procedure to maintain consistency. We have chosen to use Hodrick-Prescott (HP) filtering to separate our data into a trend and a cyclical component after carefully researching a sequence of potential filtering methods.[21] The HP-filter not only exhibits the advantage of being well known in economics, it is also the only filter separating the series into only two components. All other decompositions split the sample into at least three components, and we would therefore have to ignore the higher frequencies from our analysis. The variance of the time series is obtained from

$$\hat{\sigma}^2_{\text{HP}} = \frac{1}{m-1} \sum_{t=1}^{m} (g_t - \hat{\mu}_t)^2, \tag{24}$$

[21] See the appendix for a full discussion.

where $\hat{\mu}_t$ is the Hodrick-Prescott filtered growth rate that is obtained by solving

$$\min_{\hat{\mu}_t} \left[\sum_{t=1}^{T} (g_t - \hat{\mu}_t)^2 + \lambda \sum_{t=2}^{T-1} \left((1-L)^2 \hat{\mu}_{t+1} \right)^2 \right] \tag{25}$$

where $L^n y_t = y_{t-n} \quad \forall n \in \mathbb{N}$. The objective was to set the smoothing parameter such that for both types of stationarity, the filtered series would be a straight line. In case of non-stationarity, the filtered series should display the possible non-linear deterministic trend. Visual inspection (see figure A.3 to A.6 in the appendix suggested setting the smoothing parameter, λ, to 5000. The outcome is in line with our unit-root tests from the previous section.[22] England and the United States are stationary cases par excellence: the growth rate fluctuates around a constant value. Italy is a perfect case of trend-stationarity: the average growth rate has been declining since 1960 at a constant rate. Greece belongs in the nonstationary category: the trend growth rate was declining until the mid-1980s when it reached the bottom and started to increase again.

7.5 Results

Estimating the volatility and the average growth rate over the whole sample and running a cross-country regression afterwards would imply that we assume that both statistics are more or less stable. Visual inspection tells us that this is clearly not the case. Dividing the samples into sub-samples mitigated the effect of assuming constant volatility and constant trend growth rates. Furthermore, we end up with more data points. Of course, there is an upper-bound to the number of sub-samples since we still need enough data points to obtain a 'satisfactory' estimate of the variance (equation 24). Since the length of our time series is 45 (1961-2005) we decided to separate them into three (non-overlapping) sub-samples of length 15.[23] The resulting $3 * 21 = 63$ data points were pooled for our regression analysis.[24]

We are interested in the functional relationship between the growth rate of GDP, y, and our measure of its volatility, x. In a parametric approach, the obvious choice is linear,

$$y = \alpha + \beta x \tag{26}$$

We find a positive and significant relationship between the standard deviation and the growth rate of output,

$$y = \underset{(0.37)}{1.47} + \underset{(0.15)}{0.54}\, x \tag{27}$$

[22]Note that we have selected a λ very different from what can be found in the real business cycle literature. However, our objective, too, is very different. Whereas the real business cycle theory tries to eliminate very low frequencies (noise9), our ambition is very different: we try to split the GDP series into a trend and cyclical component.

[23]One robustness test we perform in the next chapter is splitting the sample into 2 or 4 groups. This does not alter our main findings.

[24]Pooled estimators impose the realistic assumption on our data set that the relationship between regressand and regressor is the same irrespective of whether we are looking across countries or over time within in a country, and that all the errors are drawn from the same distribution.

	$\hat{\alpha}$	**s.e.**	$\hat{\beta}$	**s.e.**
2 periods	1.49	0.46	0.53	0.18
3 periods	1.47	0.37	0.54	0.15
4 periods	1.96	0.33	0.33	0.13

Table 2: Regression estimates for different sample length

	$\hat{\alpha}$	**s.e.**	$\hat{\beta}$	**s.e.**
15-15-15	1.49	0.46	0.53	0.18
(6)-11-11-11-(6)	-1.48	0.61	0.24	0.04
(8)-15-15-(7)	-0.44	0.68	0.16	0.04

Table 3: Regression estimates for different sample length, omitting initial and final observations

where the number below the estimated coefficient indicate the standard error of the ordinary least square estimation. The regression can explain 17.7% of the variation, which is good, considering the fact that we did not include any other control variables and that we use cross country data. The result is certainly encouraging, as we find a significant relationship between economic growth and volatility. In order to confirm our results, we will conduct a series of robustness checks in the following chapter.

8 Robustness Analysis

8.1 Sample Variations

The first robustness check was to split the sample in different length. Whereas in the previous chapter, we used have split the sample in three, with a length of a single observation being 15 years, and a total of 63 observations, we have also split the sample period into 2 and 4 groups. This leads to the length of a single observation of 22 or 11 years respectively, with 42 or 84 observations. Our findings are summarized in table 2.

We obtain similar coefficient estimates for the 2-period split and the 3-period split, indicating robustness of our results. The coefficient remains statistically significant at the 5% level. The reason for the lower value may be due to the fact that 11 periods may be too short to compute the variance, and some variance is captured by the growth rates, which alter over the 4 observation periods.

The standard HP Filter is known to have problems detrending at the beginning and end of the sample period. For that reason, we created two additional series where we have eliminated the first 5 and 7 years respectively, and than split the remaining sample in three 11 year periods and two 15 year periods, respectively. The estimation results are presented in 3, and differ little from our previous results, continuing to show a positive and significant relation between economic growth and volatility.

	$\hat{\alpha}$	**s.e.**	$\hat{\beta}$	**s.e.**
Lin-Lin (26)	1.5	0.4	0.55	0.15
Log-Log (29)	1.7	1.1	0.46	0.13
Log-Lin (30)	0.5	0.1	0.17	0.05
Lin-Log (31)	1.6	0.3	1.40	0.37

Table 4: Regression estimates

8.2 Variants of Ordinary Least Squares

Regression analysis is concerned with the question of how y can be explained by x. This means a relation of the form

$$\begin{aligned} y_i &= m(x_i) + \epsilon_i \\ \mathbb{E}[Y|X=x] &= m(x). \end{aligned} \tag{28}$$

where m is a function in the mathematical sense. It determines how the *average* value of y changes as x changes. In a parametric approach, the obvious choice is linear, as discussed in the previous section, and functions whose parameters can be estimated by ordinary least squares after applying a linearizing transformation on the variables, like

$$m(x) = \alpha x^{\beta} \tag{29}$$

$$m(x) = e^{\alpha + \beta x} \tag{30}$$

$$m(x) = \alpha + \beta \ln x \tag{31}$$

In equation 29, β measures the elasticity[25] of $m(x)$ with respect to x. It can be written as $\ln m(x) = \ln \alpha + \beta \ln x$. In equation 30 β gives the proportionate change in $m(x)$ per unit change in x. Vice versa for equation 31.

Table 4 summarizes the estimation results. All four models can account for about the same amount of variability in the growth rate (between 15 and 20 percent), with the lin-lin model (26, bold solid line) and the lin-log model (31, solid line) coming out leading (see figure 1). In both models the estimate for β is significantly different from zero (p-value < 0.001). The log-log model (29, dashed line) and the lin-log model (30, dot-dashed line) still exhibit coefficient that are significant at the 5% significance level.

The coefficients cannot be compared directly, so figure 1 draws the regression lines for all four models, showing that are all very similar in the relevant area, so that we can confirm the result of the previous chapter.

So far, we have based our regression on the standard deviation as a measure of volatility. Evidently, the variance, the square of the standard deviation, may also be an indicator of volatility. Although the coefficient in equation 29, which is far from 2, suggest otherwise, we run various polynomial regressions of the more general form

[25] The elasticity measures the percent change in $m(x)$ for a 1 percent change in x. $m(x)_{\epsilon} = \frac{m'(x)x}{m(x)} = \frac{d \ln m(x)}{d \ln x}$

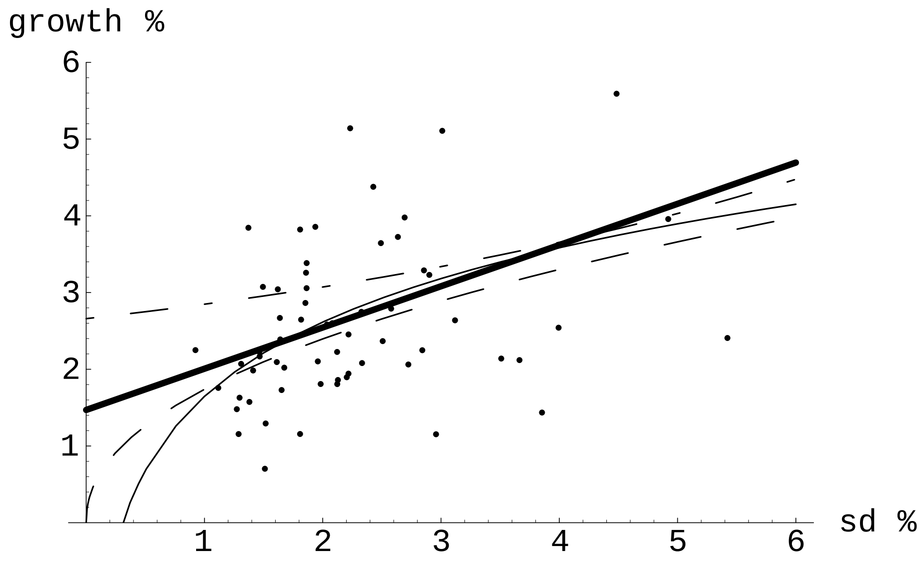

Figure 1: Scatterplot and Regression Lines

	$\hat{\alpha}$	**s.e.**	$\hat{\beta}$	**s.e.**	$\hat{\gamma}$	**s.e.**	R^2
Model 1	1.47	0.37	0.54	0.15			17.7
Model 2	2.19	0.21			8.26	2.52	15.0
Model 3	0.63	0.93	1.21	0.71	−11.57	11.77	19.0

Table 5: Regression estimates

$$m(x) = \alpha + \beta x + \gamma x^2 \tag{32}$$

We have tried higher order polynomials with no avail. The results for the estimation are presented in table 5. Whilst single variable models all yield statistically significant coefficient on the various measures of volatility, more complex models fail in obtaining these coefficients, probably due to correlation between independent variables. Among the first three models, we find that the version using the variance has a slightly higher explanatory power than the model which is based on the standard deviation, and hence preferable.

8.3 Robust Regression: M-Estimation

A statistical procedure is regarded as 'robust' if it performs reasonably well even when the assumption of the statistical model are not true. M-regression, the most common general method of robust regression introduced by Huber (1964), was specifically developed to be robust with respect to the assumption of normality (see Birkes & Dodge (1993)). Consider our linear model

$$y_i = \boldsymbol{x_i'}\boldsymbol{\beta} + \epsilon_i \tag{33}$$

for the ith of n observations. The fitted model is

$$y_i = \boldsymbol{x_i'b} + e_i \tag{34}$$

The general M-estimator minimizes the *objective function*

$$\sum_{i=1}^{n} \rho\left(e_i\right) = \sum_{i=1}^{n} \rho\left(y_i - \boldsymbol{x_i'b}\right) \tag{35}$$

where the function ρ gives the contribution of each residual to the objective function. Obviously, for least-squares estimation, $\rho\left(e_i\right) = e_i^2$. The Huber M-estimator uses a function ρ that is a compromise between e^2 and $|e|$:

$$\rho\left(e\right) = \begin{cases} e^2 & \text{for } |e| \leq \text{k} \\ 2k|e| - k^2 & \text{otherwise} \end{cases}$$

Tukey's biweight estimator is defined as:

$$\rho\left(e\right) = \begin{cases} \frac{k^2}{6}\left\{1 - \left[1 - \left(\frac{e}{k}\right)^2\right]^3\right\} & \text{for } |e| \leq \text{k} \\ \frac{k^2}{6} & \text{otherwise} \end{cases}$$

The value k for the Huber-M and Tukey's biweight estimator is called a tuning constant; smaller values of k produce more resistance to outliers, but at the expense of lower efficiency when the errors are normally distributed. We choose the pre-selected values of $k = 1.345\sigma$ for Huber's and $k = 4.685\sigma$ for Tukey's estimator (where σ is the standard deviation of the errors).

Figure 2 shows the regression lines for the OLS (red), Huber (blue), and Tukey (green) estimates. Both the Huber and the Tukey estimates of the slope are slightly lower than the OLS estimate, viz. 0.45 and 0.4, respectively, but still significantly different from zero. We can therefore still confirm the robustness of the OLS estimator presented in the previous chapter.

8.4 Detection of Influential Data Points

The purpose of any sample is to represent a certain population, actual or hypothetical. Influential data points or outliers[26] in a sample are likely to influence the sample-based estimates of the regression coefficients. There are many sources of outliers such as sampling a member not of that population, bad recording or measurement, errors in data entry, etc. For whatever reason they have come to exist, outliers will lessen the ability of the sample statistics to represent the population of interest. A common method of dealing with apparent outliers in a regression situation is to remove the outliers and then refit the regression line to the remaining points.

Since no data points that obviously qualify as an outlier could be found by visual inspection, we calculated Cook's distance for each observation. The 100(1-α)% joint confidence region for the parameter vector β is

$$\left(\hat{\beta} - \beta\right)' \left(X'X\right) \left(\hat{\beta} - \beta\right) \leq k\hat{\sigma}^2 F_{k,N-k,\alpha} \tag{36}$$

[26] Hawkins (1980) described an outlier as an observation that 'deviates so much from other observations as to arouse suspicions that it was generated by a different mechanism'. Outliers have also been labeled as contaminants (Wainer (1976))

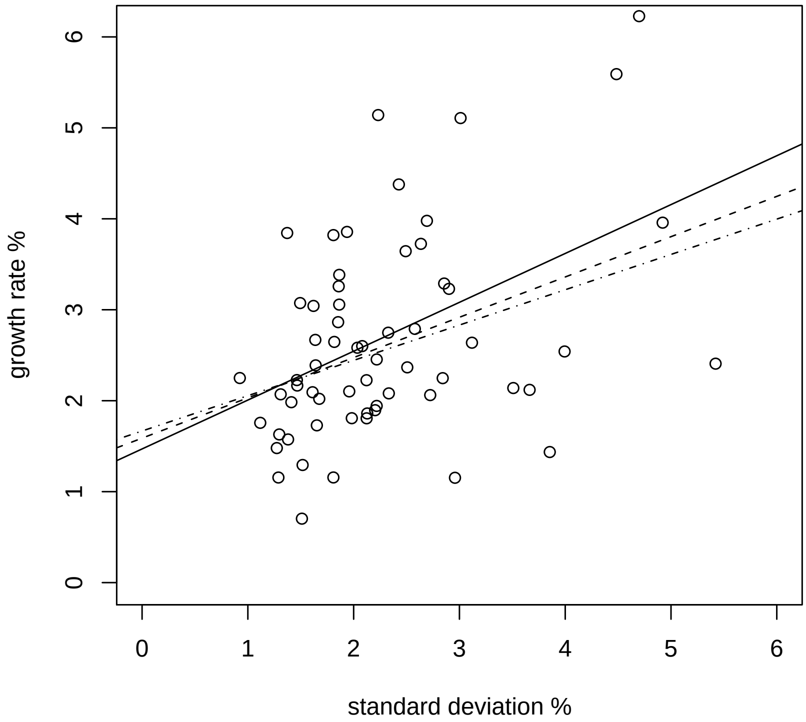

Figure 2: OLS, Huber-M, and Tukey's Biweight

Cook's Distance is defined as

$$C_i = \frac{\left(\hat{\beta} - \hat{\beta}_{-i}\right)' (X'X) \left(\hat{\beta} - \hat{\beta}_{-i}\right)}{k\hat{\sigma}^2} \tag{37}$$

The 100(1-α)% joint ellipsoidal confidence region for β given in 36 is centered at $\hat{\beta}$. The quantity C_i measures the change in the center of this ellipsoid when the ith observation is omitted, and thereby assesses its influence. C_i is the scaled distance between $\hat{\beta}$ and $\hat{\beta}_{-i}$. An alternate form of Cook's distance is

$$C_i = \frac{1}{k} \frac{h_{ii}}{(1 - h_{ii})} r_i^2 \tag{38}$$

where h_{ii} is the leverage[27] and r_i the studentized residual[28] C_is that are above the threshold value of the $50th$ percentile of the F distribution with k and

[27]The leverage assesses how far away a value of the explanatory variable is from the mean value: the farther away the observation the more leverage it has. h_{ii} is the ith diagonal element of $X(X'X)^{-1}X'$. In the bivariate case $h_{ii} = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{(n-1)s_x^2}$.

[28]The studentized residual is $r_i = \frac{e_i}{s_e\sqrt{1-h_{ii}}}$.

N-k degrees of freedom (in our case 0.7) are regarded as influential observations. According to this definition, as can be seen in 3, our sample does not contain any influential observations.

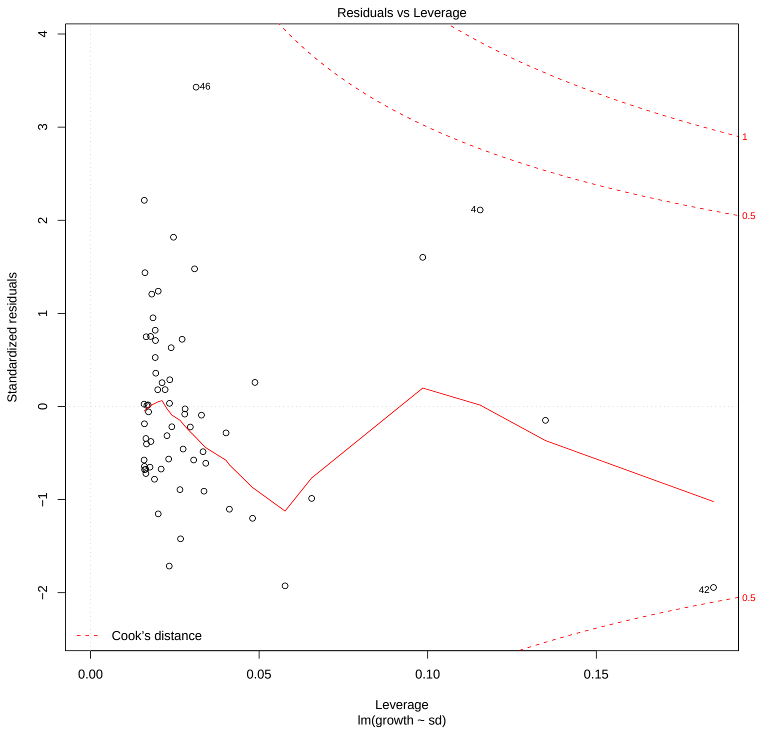

Figure 3: Influential Data Points

The most influential data points in our sample are Greece$_{1960-75}$ (#4) with a growth rate of 6.2% and a standard deviation of 4.7%, Turkey$_{1990-05}$ (# 42) with a growth rate of 2.4% and a standard deviation of 5.4%, and Japan$_{1960-75}$ (#46) with a growth rate of 7% and a sd of 3.2%. Running a OLS regression without those three data points yielded a slope of 0.46, wich is the same result as the one obtained by using the Huber-M-Estimator. Once again, this confirms our results of a positive and significant relationship between economic growth and volatility.

8.5 Nonparametric Estimation: Kernel Regression

Our final test of robustness is to use nonparametric estimation methods. The nonparametric approach does not assume any functional form for $m(x)$, but rather goes back to the statistical definition of conditional expectation:

$$m(x) = \mathbb{E}[Y|X=x] = \int_{-\infty}^{+\infty} y f_{Y|X}(y|x)\, dy = \frac{1}{f_X(x)} \int_{-\infty}^{+\infty} y f_{X,Y}(x,y)\, dy \tag{39}$$

Plugging in Kernel estimates for the marginal density, $f_X(x)$, and the joint density, $f_{Y,X}(y,x)$, delivers an estimate $m(x)$ of the conditional expectation at point x:

$$\frac{1}{\hat{f}_X(x)} \int_{-\infty}^{+\infty} y \hat{f}_{X,Y}(x,y)\, dy \tag{40}$$

This has become known as the Nadaraya-Watson estimator. Figure 4 shows two Nadaraya-Watson regression estimates, one with high bandwidth (dark blue line) and one with low bandwidth (light blue line). In the dense region, i.e. in the region where many data points are available, the estimates tell the same story as the OLS regression line, so it seems that there really is a linear relationship between volatility and growth. The Nadaraya-Watson estimates become very erratic in the region where the standard deviation is larger than 3.5%. This was to be expected, since only eight data points fall into this region.

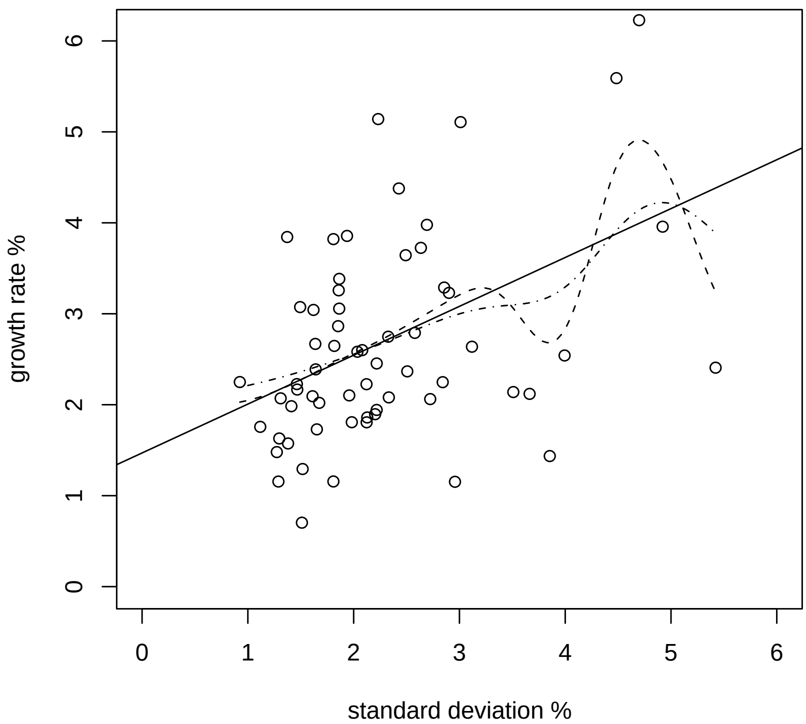

Figure 4: Nadaraya-Watson Estimates and OLS Regression Line

After running an entire series of robustness tests, from altering the sample, running non-linear versions of OLS regressions, M-estimations, checking against critical data points, and nonparametric methods, which all point toward a positive and significant relationship between economic growth and volatility, we are convinced about the robustness of our results indicated in the previous chapter.

9 Conclusion

The contribution of this paper is twofold. First the empirical result of a robust and positive relationship between economic growth and volatility should stimulate and support further theoretical research in the field, which is growing in magnitude and importance. Second, the paper suggests an empirical method to analyze the relationship between economic growth and volatility. We use the well-known Hodrick-Prescott filter to separate GDP time series into a trend component and a cyclical component, and then use period averages to obtain statistics for growth and volatility. This method is preferential to other band-pass filtering techniques, but also with respect to GARCH methods, which are wholly unfit for short time series such as national accounting data.

Using the time series experience of twenty-one OECD countries between 1961 and 2005, we have presented strong empirical evidence for a positive relationship between output variability and economic growth. This relationship is robust against outliers and also shows up in a non-parametric setting. A case can be made that our measure of output variability is more suitable than the ones used in previous work for time series of economic growth.

These results have to be treated with care, particularly when making policy implications. Whilst we find that there is a positive and significant relation between economic growth and volatility, we refrain from making any comment on causality. Factors that increase volatility, such a pro-cyclical fiscal or monetary policy probably will not alter the growth pattern of the economy. We do believe in "innovative risk", or the concept that an innovation, which will induce economic growth, is intrinsically risky, and therefore we should observe a positive relation between growth and volatility in the data, as we indeed do. Whilst it is true, at least at the margin, that an increase in innovation would lead to faster economic growth, this will come at the cost of higher volatility. We think that economic stability is welfare enhancing, and therefore policymakers face a trade-off between economic growth and volatility.

References

Aghion, P., Angeletos, G.-M., Banerjee, A. & Manova, K. (2005), Volatility and growth: Credit constraints and productivity-enhancing investment, Technical report, Harvard University.

Aghion, P. & Saint-Paul, G. (1993), Uncovering some causal relationships between productivity growth and the structure of economic fluctuations: A tentative survey, Technical Report 4603, National Bureau of Economic Research, Inc.

Baxter, M. & King, R. G. (1995), Measuring business cylces: Approximate band-pass filters for economic time series, Technical Report 5022, National Bureau of Economic Research, Inc.

Bernanke, B. (1983), 'Irreversibility, uncertainty, and cyclical investment', *Quarterly Journal of Economics* **98**(1), 85–106.

Birkes, D. & Dodge, Y. (1993), *Alternative Methods of Regression*, John Wiley & Sons, New York.

Black, F. (1981), 'The abcs of business cycles', *Financial Analysts Journal* **37**(6), 75–80.

Boulding, K. E. (1966), *Economic Analysis, Volume I: Microeconomics*, fourth edn, Harper & Row, New York.

Campbell, J. & Mankiw, N. G. (1987), 'Are output fluctuations persistent?', *Quarterly Journal of Economics* **102**(4), 857–880.

Campbell, J. & Mankiw, N. G. (1989), 'International evidence on the persistence of economic fluctuations', *Journal of Monetary Economics* **23**(2), 319–333.

Caporale, T. & McKiernan, B. (1998), 'The fischer black hypothesis: Some time-series evidence', *Southern Economic Journal* **64**(3), 765–771.

Grier, K. B. & Perry, M. J. (2000), 'The effects of real and nominal uncertainty on inflation and output growth: Some garch-m evidence', *Journal of Applied Econometrics* **15**(1), 45–58.

Grier, K. & Tullock, G. (1989), 'An empirical analysis of cross-national economic growth, 1951-80', *Journal of Monetary Economics* **24**(2), 259–76.

Hawkins, D. (1980), *Identification of Outliers*, Chapman and Hall, London.

Henderson, R. (1924), 'A new method of graduation', *Transactions of the Actuarial Society of America* **25**, 29–40.

Huber, P. (1964), 'Robust estimation of a location parameter', *Annals of Mathematical Statistics* **35**, 73–101.

Jovanovich, B. (2006), 'Asymmetric cycles', *Review of Economic Studies* **73**(1), 145 –162.

Kormendi, R. & Mequire, P. (1985), 'Macroeconomic determinants of growth: Cross-country evidence', *Journal of Monetary Economics* **16**(2), 1345–70.

Kwiatkowski, D., Phillips, P., Schmidt, P. & Shin, Y. (1992), 'Testing the null hypothesis of stationarity against the alternative of a unit root : How sure are we that economic time series have a unit root?', *Journal of Econometrics* **54**(1-3), 159–178.

Leland, H. E. (1968), 'Saving and uncertainty: The precautionary demand for saving', *The Quarterly Journal of Economics* **82**(3), 465–473.

Leser, C. (1961), 'A simple method of trend construction', *Journal of the Royal Statistical Society. Series B (Methodological)* **23**, 91–107.

Levine, R. & Renelt, D. (1992), 'A sensitivity analysis of cross-country growth regressions', *American Economic Review* **82**(4), 942–63.

Mills, T. (2000), 'Business cycle volatility and economic growth: A reassessment', *Journal of Post-Keynesian Economics* **23**, 107–116.

Nelson, D. (1990), 'Stationarity and persistence in the garch(1,1) model', *Econometric Theory* **6**(3), 318–34.

Phillips, P. & Solo, V. (1989), Asymptotics for linear processes, Technical Report 932, Cowles Foundation.

Pindyck, R. S. (1991), 'Irreversibility, uncertainty, and investment', *Journal of Economic Literature* **29**(3), 1110–48.

Ramey, G. & Ramey, V. A. (1995), 'Cross-country evidence on the link between volatility and growth', *American Economic Review* **85**(5), 1138–51.

Sandmo, A. (1970), 'The effect of uncertainty on saving decisions', *Review of Economic Studies* **37**(3), 353–360.

Schumpeter, J. A. (1939), *Business Cycles*, first edn, McGraw-Hill, New York.

Tsay, R. (2005), *Analysis of Financial Time Series*, second edn, John Wiley & Sons, Inc, Hoboken, New Jersey.

Wainer, H. (1976), 'Robust statistics: A survey and some prescriptions', *Journal of Educational Statistics* **1**(4), 285–312.

Whittaker, E. T. (1923), 'On a new method of graduation', *Proceedings of the Edinburgh Mathematical Society* **41**, 63–75.

Zarnowitz, V. (1981), Business cycles and growth: Some reflections and measures, *in* W. Muckl & A. Ott, eds, 'Wirtschaftstheorie und Wirtschaftspolitik: Gedenkschrift für Erich Preiser', Passavia Universitätsverlag, Passau, pp. 475–508.

APPENDIX

A Garch-in-Mean Regression Models

In the GARCH-in-Mean (GARCH-M) model the conditional variance of the error term is used as an explanatory variable in the equation (19) for the conditional mean of the variable to be explained. The error term follows a GARCH(p,q) model

$$u_t = \sigma_t \epsilon_t \tag{A.1}$$

where $\epsilon_t \sim IID(0,1)$ and σ_t^2, the conditional variance of u_t conditional on all the information up to time $t-1$, $\mathcal{F}_{t-1}$, is given as:

$$\mathbb{E}\left[u_t^2|\mathcal{F}_{t-1}\right] = \sigma_t^2 = \omega + \sum_{j=1}^{q} \alpha_j u_{t-j}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{A.2}$$

All coefficients in equation A.2 are necessarily non-negative. Nelson (1990) showed that a GARCH(1,1) process is strictly stationary when $\mathbb{E}[\log(\alpha\epsilon_t^2 + \beta)] < 0$. When $\epsilon_t \sim N(0,1)$, the condition for strict stationarity is weaker then the condition for covariance stationarity $\alpha + \beta < 1$.

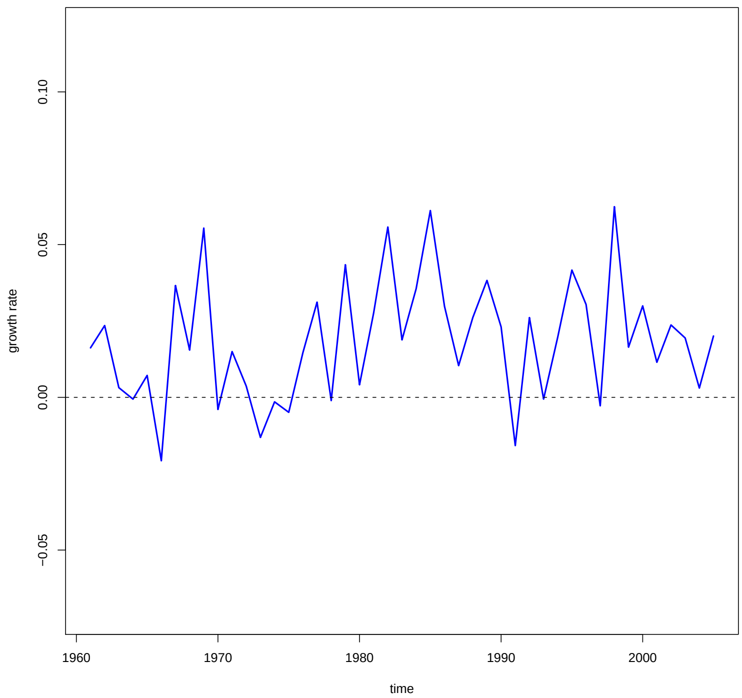

Figure A.1: Trajectory of a GARCH(1,1)-M process

Figure A.1 shows a trajectory of a GARCH(1,1)-M process. The risk premium parameter, γ, was set to 2, a value in between those obtained by the

GARCH(0,1)-M model of Caporale et al. (0.7) and the bivariate GARCH(1,1)-M model Grier et al. (3.5). The parameters for the variance equation, α and β, were set to 0.1 and 0.8, respectively. These values are common in finance (see for instance Tsay (2005)) and close to the ones obtained by Grier & Perry (2000) (0.2 and 0.7).[29] Though it seems that such processes are capable of producing series that resemble actual GDP growth rates, unfortunately, very long time series (n » 2500) are required for estimating such processes efficiently.

In a small Monte-Carlo simulation running 100 realizations of a GARCH(1,1)-M process with t = 1,...,200 and with the parameters as given above and re-estimating the process yielded the distribution of the GARCH-in-Mean effect, $\hat{\gamma}$ as shown in figure A.2.

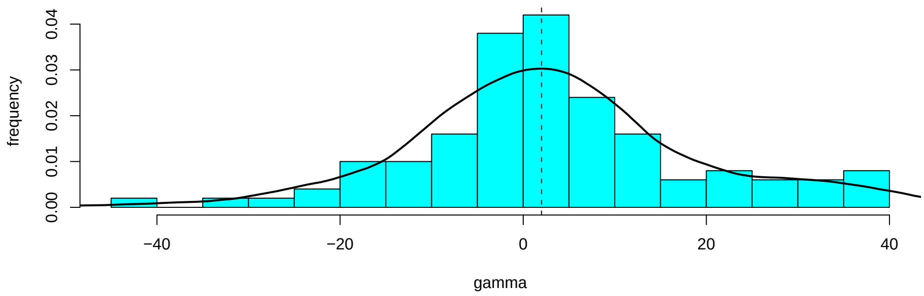

Figure A.2: Histogram and Empirical Density Function

The average is close to the true mean of our simulation (3 instead of 2) but the standard deviation of 15 is unacceptably large. In 25 percent of our simulation we obtained an estimate for γ that was at least twice as large but had the opposite sign (-4 instead of 2). Apart from this technical obstacle, the implication of the fact that the measure for volatility is based solely on forecast uncertainty seems to be not fully understood when the mean equation 19 contains additional regressors.

B Growth Rates and the HP filter

[29] The intercepts were set to $\omega = 0.0001$ and $\kappa = 0.005$, respectively and $\epsilon \sim \mathcal{N}(0,1)$

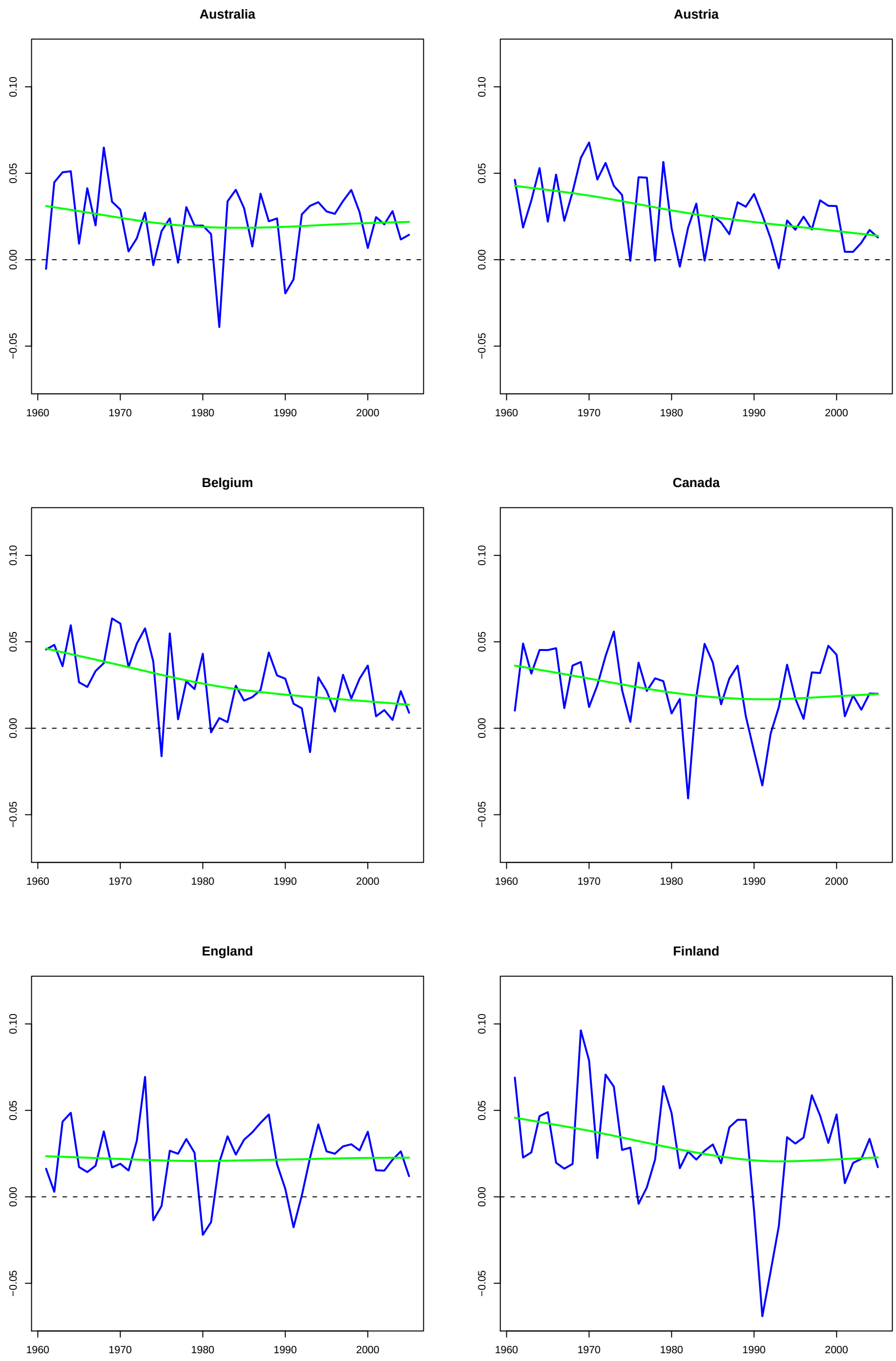

Figure A.3: Growth Rates of Selected Countries 1

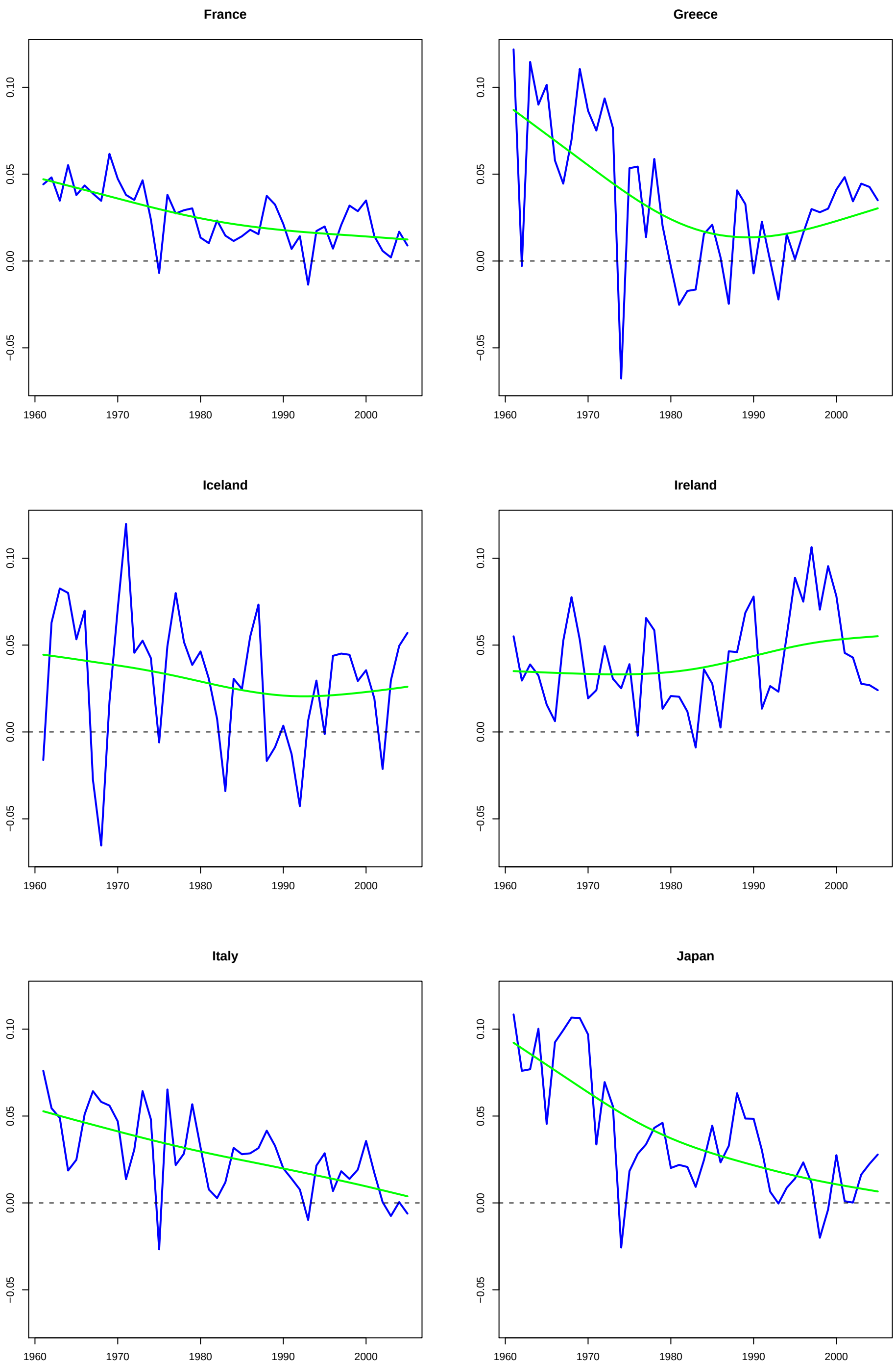

Figure A.4: Growth Rates of Selected Countries 2

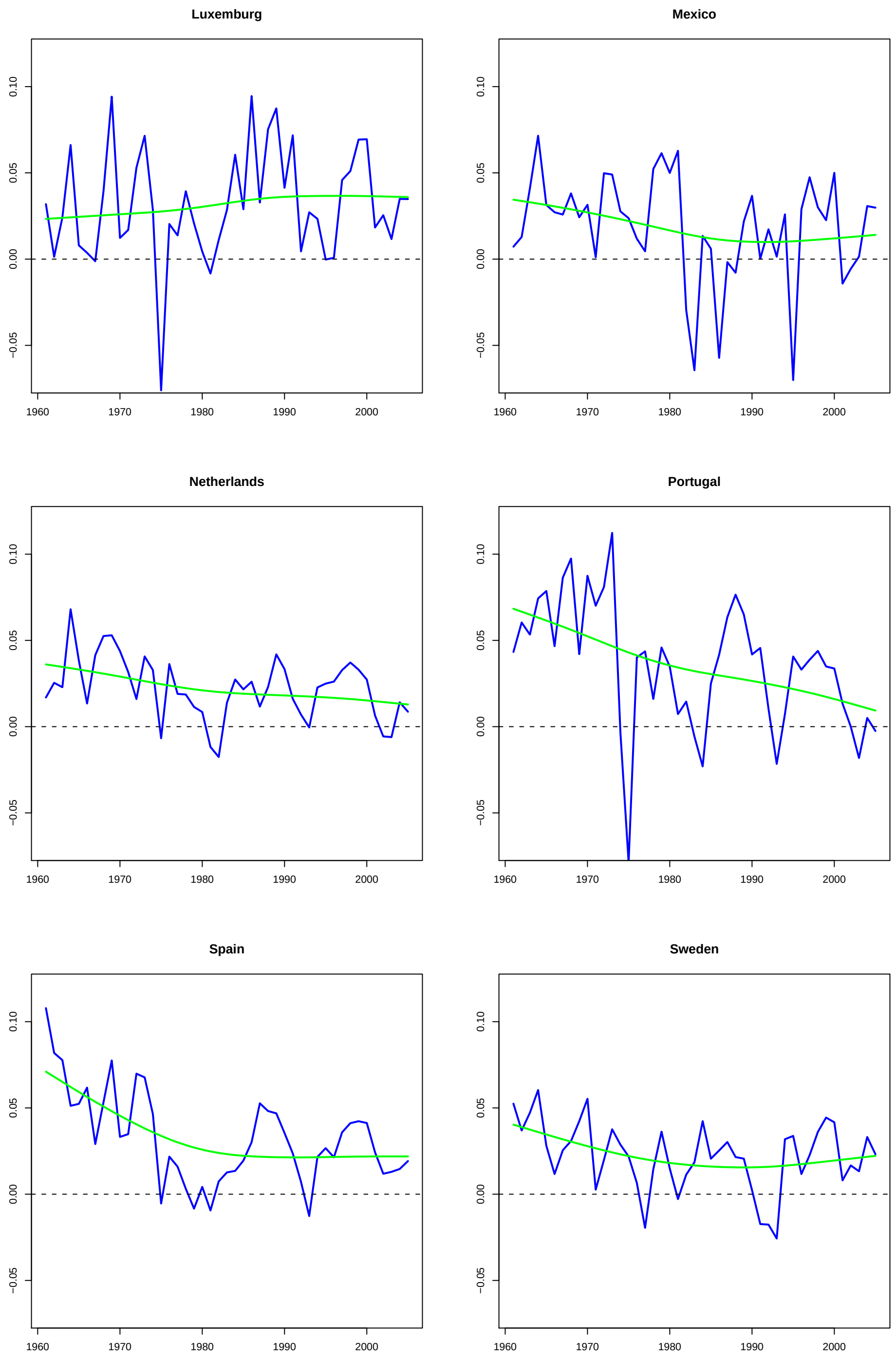

Figure A.5: Growth Rates of Selected Countries 3

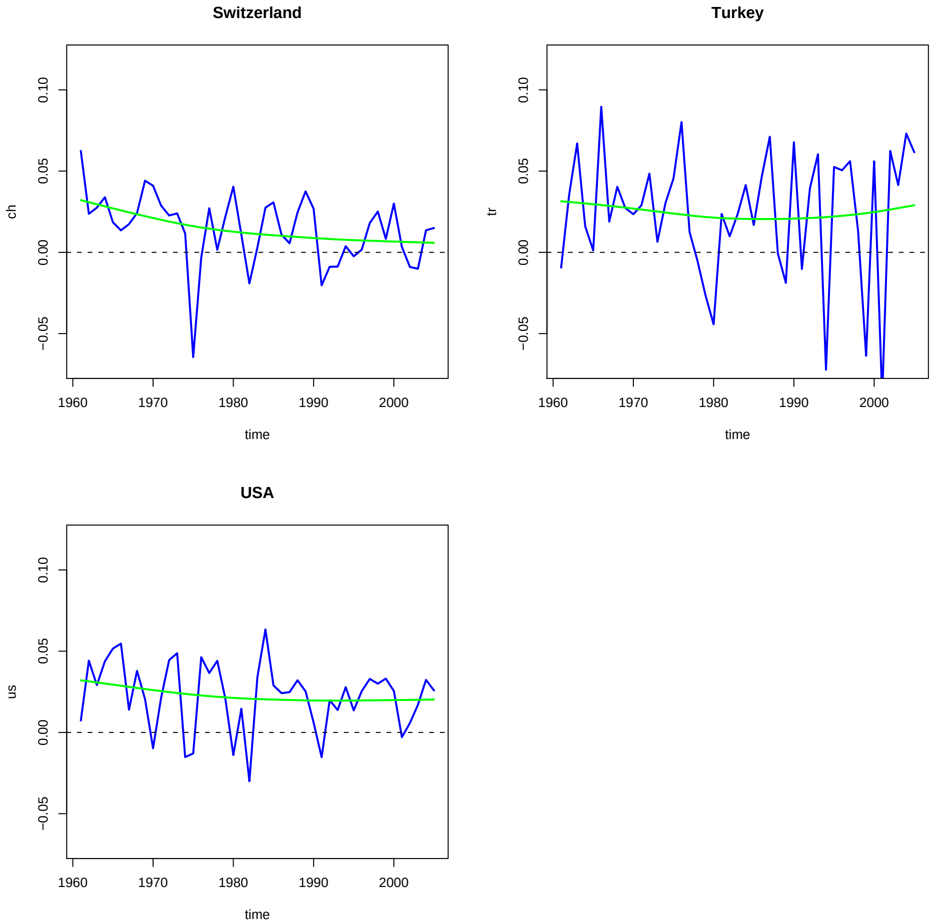

Figure A.6: Growth Rates of Selected Countries 4