Bachelor Thesis

The View of the Optimist: Does Entrepreneurship predict the Phase of the Business Cycle?

Tim Lamballais Tessensohn
321789 TL

Supervisors
Prof. Dr. A.R. Thurik
Prof. Dr. Ir. R. Dekker
Dr. W. van den Heuvel

17:00, 01-07-2011, Rotterdam

Econometrics and Operations Research
Erasmus University Rotterdam

Table of Contents

1. Introduction

Both entrepreneurship and business cycles have enjoyed a growing interest by researchers and politicians. Entrepreneurship in modern economies was falling before the seventies but has since recovered[1] and is perceived to be an important factor in economic growth, especially in high-tech and high-growth industries. Regarding economic growth, the economic recession of the past few years has sparked a renewed concern about business cycles and how to influence their dynamics.

The question rises what kind of relationship there is between entrepreneurship and the business cycle. Entrepreneurship is to some degree susceptible to the influences of national entrepreneurship policies. Is it possible for politicians to influence their national business cycle through entrepreneurship policies?

There have been some theoretical arguments and models regarding the role of entrepreneurship in causing business cycle fluctuations but there appears to have only been one empirical and international research, which was done by Koellinger and Thurik (2012). They discern two macroeconomic levels, an aggregate, global level and a national level. The variables on the aggregate, global level are weighted sums of the variables on the national level. They find that on the aggregate level entrepreneurship is pro-cyclical and granger-causes the business cycle. GDP and unemployment do not predict self-employment. These results differ for the national level, where the relationship between self-employment and entrepreneurship is less clear and unemployment affects self-employment.

The weaker results for the national level may be due to the influence of the world cycle on the national business cycles. As countries are connected through trade it would seem plausible that national business cycles do not excessively deviate from their group of major trading partners. This synchronizing effect may obscure solid relationships because it complicates causal relationships. Another reason would be country specific shocks, such as national economic policies. These may blur and disturb a direct relation between self-employment and the business cycle.

Hence, the underlying dynamics may be better understood from a perspective with less detail. Being able to predict more or less exactly where the cycle will be in the future may not be realistic at the national level but predictions about the phases of the cycle can be. The approach is therefore ‘qualitative’ rather than ‘quantitative’. This thesis builds on the ‘quantitative approach’ of Koellinger and Thurik (2012) and attempts to complement it. It aims to examine whether entrepreneurship on the national level forecasts booms and recessions. Perhaps his optimistic view allows the entrepreneur to predict a transition from recessions to booms?

This thesis will follow the ‘let the data speak freely’ approach of Koellinger and Thurik. As their results are at odds with theoretical frameworks[2], a 'reality first' approach seems more sensible.[3] This means that no new

[1] Wennekers et al (2010)
[2] These would be the theoretical frameworks of Rampini (2004) and Bernanke and Gertler (1989), see also the section 'Literature review'.
[3] For a discussion about advantages and disadvantages of the 'theory first' as opposed to the 'reality first' approach, see Juselius (2009).

hypothesis will be formulated nor will incumbent hypothesis be tested. It also means that there will be no theoretical framework from which the results are interpreted.

This thesis is structured as follows: section 2 defines the problem research questions and objective. Section 3 reviews the literature. Section 4 investigates the peculiarities of the data and data transformations. Section 5 deals with the research and modeling methods. Firstly, the relationship between certain cycle characteristics and macroeconomic variables are examined at the level of the cycle rather than on a yearly basis. Then, the efforts are towards modeling the phases and transitions in the business cycles as preparations for forecasting. The outcomes are reported in section 6 for the whole data set, whereas section 7 explores their outcomes in forecasting and sideward replication. Section 8 concludes and discusses the results.

2. Research Questions and Problem Definition

The topic of this thesis is the influence of entrepreneurship on national business cycles. The problem definition would be:

"How do national business cycle characteristics depend on entrepreneurship?"

This main question will be answered by focusing on the following sub-questions:

1) “Which component(s) of business cycles is/are significantly influenced by entrepreneurship – length, amplitude, slope or mass?”
2) “To what extent do the global business cycle, global and national self-employment and national unemployment influence which phase the national business cycle passes through?”
3) “Do the influences mentioned at 2) change for different phases of the national business cycle?”
4) “Are the variables mentioned at 2) able to forecast the business cycle?”

Cycle characteristics would be the length of a cycle, the maximum height (amplitude) the cycle reaches, the average slope and the 'mass'. Mass here is defined as the integral of the cycle (the 'surface between the time series and x-axis'). Besides these there are also four phases and two transitions: at any point the business cycle is either in a boom or recession (positive or negative cyclical component) and has either a positive slope or a negative slope. The phases are therefore categorized into boom-positive slope (BP), boom-negative slope (BN), recession-negative slope (RN) and recession-positive slope (RP). It is not possible to move from RN or RP to BN and it is also not possible to move from BP or BN to RP. Furthermore there are two transitions, from boom to recession (TBR) and from recession to boom (TRB).

A complicating factor would be interactions between countries. As countries trade intensively with one another, assuming independence of business cycles across countries would turn this thesis into mere fireplace-fodder. It stands to reason that the national business cycle cannot run too much out of synch with that of their major trading partners. Unfortunately, due to the lack of export and import figures in the countries and time period covered it is impossible to measure trade relationships across countries. Instead there will be a 'world business cycle', an aggregated global business cycle created from summing over the national business cycles. Similarly, there will be global, aggregated self-employment and unemployment. These three global variables appear as explanatory variables in the regressions. The consequence is that there is no strict separation between the global and the national level as is the case of Koellinger and Thurik (2012).

The objective is to recommend whether a politician should stimulate national entrepreneurship policy to create more beneficial business cycle behavior, given world influences on this behavior. Here beneficial would be shorter or less severe recessions and longer booms periods. The predicted effect will depend on the characteristics of the phases of the business cycle, whether it is in a recession or in a boom and whether it is increasing or declining. Stimulating entrepreneurship may increase the tendency to remain in a boom. A longer boom length may however come at the cost of lower amplitude, which justifies examining cycle characteristics beside the cycle phases. This approach does not seem to have been followed before.

Note again that in this thesis the 'data is allowed to speak freely' without the constraints of a theoretical model. This also means that a range of different techniques will be used. OLS, Poisson and Negative Binomial regressions are used to look at cycle components such as length, amplitude and mass. Binary and multinomial logit models to examine the phases and transitions.

3. Literature Review

Traditionally, business cycles have been analyzed with factor models, incorporating capital, labor and exogenous technology shocks (Kydland and Prescott 1982). Such technology shocks have been linked to entrepreneurship already by Schumpeter (1934) as entrepreneurs may challenge incumbent firms with innovations that may render older technology obsolete. Corriveau (1994) presents a theoretical model where these technology shocks are endogenously caused by entrepreneurs.

Innovations are sometimes thought to occur in waves, which prompted the birth of theoretical economic models linking 'implementation cycles' and entrepreneurship to the business cycle (Shleifer 1986, Francois and Lloyd-Ellis 2008). Complementary to such models would be the Caballero and Hammour (1991) model where recessions are utilized as a 'cleansing period' where outdated technology is 'pruned out'. These models revolve around the perception that technology shocks are of profound importance to the business cycle, an observation which is not undisputed (Eichenbaum 1991).

Other theoretical arguments linked entrepreneurship to the business cycle through credit, loans and equity. The key to this linkage is the principal-agent problem, as lenders do not have the same information as the entrepreneur, both about his skills and his products. Bernanke and Gertler (1989) introduce a model where agency costs are endogenous and depend on the entrepreneur's net-worth. This results in more investments during booms as agency costs are lower when the agent's net worth is higher. Carlstrom and Fuerst (1997) build on this by creating a computable model that is able to quantify the diffusion of productivity shocks by means of agency costs.

Bernanke and Gertler (1989) and Carlstrom and Fuerst (1997) assume the proportion of entrepreneurship in the population is constant, whereas Rampini (2004) shows a model where the number of entrepreneurs is endogenously caused by the business cycle. As agents are risk-averse and this aversion decreases with wealth increase, Rampini argues that entrepreneurship is pro-cyclical. One of his main assumptions, that entrepreneurship enjoys returns higher than wages, does not hold in practice (van Praag and Versloot 2007).

There is a vast body of empirical literature linking unemployment and the business cycle and there is research that connects unemployment and entrepreneurship (see for example Faria et al 2009). Such connections can be “refugee effects” where the unemployed start new business and “entrepreneurial effects” where increased entrepreneurship reduces unemployment later on (Thurik et al 2008). Despite these connections, the only empirical study incorporating entrepreneurship, unemployment and the business cycle seems to be Koellinger and Thurik (2012). An empirical study by Congregado et al (2009) examines the empirical relationship between the business cycle and entrepreneurship but they focus on how the business cycle influences entrepreneurship.

When it comes to methodology, time series analysis seems the obvious way forward to study the business cycle.[4] Time series models have incorporated the asymmetric dynamics between booms and recessions by utilizing Markov switch models (Kim and Nelson 1999, Filardo 1994). Birchenhall et al (1999) show that a Logistic Classification model performs superior to the Markov Switching model in predicting the states of the system. The rationale behind Markov Switching models and the nonlinear TAR models employed by Congregado et al (2009) and Potter (1995) is asymmetry in the business cycle. Asymmetry in the sense that dynamics may be different during booms then during recessions. Potter (1995) shows for instance that the effects of shock on the business cycle is different during booms and recessions. Those models included only two phases, booms and recessions.

The asymmetry may imply that entrepreneurship can be a leading indicator for (switching to) booms but not for recessions or vice versa. An explanation for this would be that entrepreneurs tend to be more optimistic than the average person.[5] Entrepreneurs may therefore be better at predicting transitions from a recession to a boom than vice versa.[6]

A truly complicating factor in empirical research on the business cycle would be the existence of chaotic determinism in the relationship between the business cycle and other macroeconomic variables. The business cycle would still have a pattern but one that is too irregular to lend itself to forecasting[7]. There do not seem to be misspecification tests for this kind of behavior, except perhaps for a runs test on randomness.

Although the objective of this thesis implies some kind of policy recommendation, the connection between government policies and entrepreneurship falls outside its scope. The interested reader is referred to Audretsch, Grilo and Thurik (2007) and Thurik (2009) for more information on the impact of policies on entrepreneurship and the channels through which this occurs.

[4]For instance, Forni and Reichlin (1998) show with a factor time series analyses how the co-movement between sectors in the United States explains part of the variance in industrial output. Replace ‘sectors’ by ‘countries’ and there is a strong resemblance with the situation in this thesis.

[5] This formed the inspiration for the title of this thesis.

[6] See Koellinger et al (2008) and Camerer and Lovallo (1999) for papers on the optimism and overconfidence of entrepreneurs.

[7] This is due to ‘sensitive dependence on initial conditions’, see Gleick (1987). The business cycle is known for its recurrent though irregular fluctuations (Burda and Wyplosz 2005) which could indicate chaotic features. The interested reader is referred to Lorenz (1987) and Puu and Sushko (2004) for theoretical explorations into chaos and the business cycle.

4. Data

The data used in this paper is the COMPENDIA dataset from EIM, where COMPENDIA stands for "COMParative ENtrepreneurship Data for International Analysis". The COMPENDIA data is a harmonized dataset covering 23 countries over the time period 1972 to 2008 and including a sizeable number of other macroeconomic variables, such as for instance GDP per capita, average number of workers per firm and labor productivity[8]. Differences in data definitions have frustrated empirical research and comparisons in the past, notably in Congregado et al (2009). Data definition differences across countries and data definition breaks in time series of individual countries have been corrected in COMPENDIA.

A complicating factor is the unification of Germany. Koellinger and Thurik (2012) opted to exclude it from their analysis because they could not correct for it. Yet reconstruction of their results[9] with and without Germany suggests that Germany's influence is rather strong. Moreover, the COMPENDIA data seems to have been corrected for the unification by using the AMECO database[10]. Germany is therefore included in the analysis. There do not seem to be other outliers in the dataset.

The concept 'entrepreneurship' is represented in this research by the 'self-employment rate', the number of business owners as a share of the labor force. Total labor force consists of employees, unpaid family workers, self-employed and unemployed persons. This operationalization does hide important aspects. Golpe (2009) and Koellinger and Thurik (2012) show that different subcategories of entrepreneurship, such as innovative and imitative entrepreneurship, interact differently with the cycle. Amalgamation of these subcategories into one self-employment rate could cause the effects of subcategories to cancel each other out. The total number of explanatory variables equals 81. This includes interaction between explanatory variables, squares, cyclical components and interaction of cyclical components of certain explanatory variables. Time dummies, country dummies, global aggregated variables and business cycle characteristics on a yearly basis complement this.

4.1 HP-Filter

The data about Gross Domestic Product (GDP) is 'raw' meaning that the cyclical component must be separated from the trend component. This can be done with a Kalman filter, as in Congregado et al (2009), though the Hodrick-Prescott filter (HP-filter) seems to be the most popular method. The method (Hodrick and Prescott 1997) distinguishes between the times series y_t, the cyclical component c_t and the growth component g_t.

$$y_t = g_t + c_t \qquad for\ t = 1, \ldots, T \qquad (1)$$

$$min_{\{g_t\}_{t=-1}^{T}} \{\textstyle\sum_{t=1}^{T} c_t^2 + \lambda \sum_{t=1}^{T}[\,(g_t - g_{t-1}) - (\,g_{t-1} - g_{t-2})]^2\} \qquad (2)$$

After selecting a value for λ, the growth or trend component can be calculated in (2) and hence from (1) also the cyclical component c_t. Once de-trended, the time series of the cyclical component c_t is divided by the original GDP time series to express the cyclical component as a percentage.

---

[8] See also van Stel (2005) for more information about this harmonization.
[9] This reconstruction is not included in this thesis as it adds little value and the appendices are cramped already.
[10] The manner in which this has happened is beyond the scope of this thesis though it seems safe to trust in the craftsmanship of Dr. André van Stel, composer of the COMPENDIA dataset.

Although a standard practice for de-trending time series, the HP-filter does not appear flawless. As Ahumada and Garegnani (2000) point out, there can be 'spurious' cases and the HP-filter may impose a cyclical structure where there is none. Moreover, the standard value of λ for yearly data, which would be λ = 100, is not without controversy, see also Appendix B.

4.2 Panel Data and Fixed and Random Effects

The data allow for the construction of a strongly balanced panel with countries as units of observation. Special about the panel structure is the possibility of controlling for 'heterogeneity': differences between the observation units that are not observed. These so-called unobserved effects may be structurally different from unit to unit and correlated with the explanatory variables. Unobserved effects usually capture the effects of variables that should be included in the regression but cannot be, for instance because the variables prove immeasurable. Culture would be an example of an unobserved effect.

The unobserved effects in panel structures allow for five approaches: pooled estimation, random effects, fixed effects, population averaging and mixed effects. In all approaches the coefficients for the explanatory variables are equal across observation units except for the coefficient of the constant term or intercept. The population averaging and mixed effects approaches were not considered in this thesis. Unobserved effects models for panel structures were not addressed in the Bachelor courses, which is why the following contains a short exploration.

In pooled estimation the coefficient for the constant term, denoted 'c', is equal across units of observation. This approach ignores heterogeneity. Units of observation always differ but perhaps not significantly so, in which case pooled estimation is a sensible option.

Random effects models the coefficient of the constant by assuming a coefficient 'c' equal across observation units and adding a 'disturbance' term 'u_i', which is a random variable, where $u_i \sim IID(0, \sigma_\alpha^2)$. This term is specific per unit i and constant throughout time. Typically the observation units are randomly selected from a larger pool. Closed form estimators exist for random effect probit models, though not for logit models (Wooldridge 2002). Disadvantages of the random effects models are that they assume a certain distribution of u_i and assume that the u_i are not correlated with the other explanatory variables. The latter disadvantage is solved by Chamberlain's random effect probit model (Chamberlain 1980). In that model the assumption changes to[11]:

$$c_i | \, x_i \sim Normal(\psi + \bar{x}_i \xi, \sigma_\alpha^2) \qquad (3)$$

(see also Mundlak 1978). Formula (3) allows for correlation between the unobserved effects and the mean of the explanatory variables through a coefficient ξ. This suggests that if there are random effects, Chamberlains random effects probit model is more appropriate than the random effects logit model.

Fixed effects dispatch the notion of a coefficient 'c' which is the same across units and replaces it with a unit specific coefficient for the constant term, 'c_i', where i denotes the unit. The estimation of all these constant terms in addition to the other coefficients results in the "incidental parameter problem" (Lancaster 2000). This

[11] See Appendix B for a more detailed explanation.

problem can be circumvented with fixed effects logit models, where the unobserved effects 'c_i' are 'conditioned away'[12], which allows for consistent, though biased, estimation of the other coefficients (Greene 2004, Katz 2001, Mundlak 1978, Wooldridge 2002). This is a quality specific to fixed effects logit models which renders it superior to fixed effect probit. It allows for correlations between the explanatory variables and the unobserved effects while simultaneously not requiring assumptions about the distribution of those unobserved effects. However, the fixed effects logit model does not allow forecasting precisely because it circumvents rather than solves the incidental parameter problem.

It is possible to test for heterogeneity. The test is an LR-test[13] which involves the test statistic $\rho = \frac{\sigma_\alpha^2}{\sigma_\alpha^2+\sigma_\epsilon^2}$ which follows an adjusted χ^2 distribution, the $\bar{\chi}^2(1)$ distribution (Andrews 1988, Self and Liang 1987, Stram and Lee 1994). Here σ_ϵ^2 is de variance of the disturbance terms and σ_α^2 is the variance of the unobserved effects. The null hypothesis is $H_0: \rho = 0$.

Beforehand it would be plausible that unobserved effects have a significant influence on the business cycle, its phases and its transitions. Unobserved effects could capture the openness of a country to trade[14], its culture or its economic policies. These should be related to other explanatory variables such as self-employment and unemployment, which means that a fixed effects model seems more suitable. Although this does not allow forecasting it can be used to obtain estimates for marginal effects and elasticities.

4.3 Randomness and Correlation in the Business Cycle Phases

Let ys be a variable denoting the phase of the business cycle. Then ys can take four values, 1 for BP, 2 for BN, 3 for RN and 4 for RP[15]. It is possible to test the randomness of the sequence ys with a runs test for each phase separately. The runs test as explained by Wackerly et al (2008) examines the randomness of a sequence of numbers that can take only two values. A run is a subsequence in that sequence whose numbers all have the same values. The number directly preceding or following the run must have a different value from the numbers in the run. Let R be the number of runs. Then the test statistic Z is given by[16]:

$$Z = \frac{R-E(R)}{\sqrt{V(R)}}, \qquad Z \sim N(0,1) \tag{4}$$

$$E(R) = \frac{2n_1n_2}{n_1+n_2} + 1 \tag{5}$$

$$V(R) = \frac{2n_1n_2(2n_1n_2 - n_1 - n_2)}{(n_1+n_2)^2(n_1+n_2-1)} \tag{6}$$

A rather large or a quite small number of runs would lead to the rejection of the null hypothesis that the sequence is random. The runs test is done both for the whole sample as well as for individual countries and both for the whole time period as well as for the time period 2005-2008 that will be forecasted. The test is performed on binary variables for each phase and transition and equal one if the business cycle is in that phase

[12] See Appendix B for a more detailed explanation.
[13] An alternative would be the Lagrange Multiplier test of Hausman, see Hausman (1978). However, the mentioned LR test is amidst the standard output of the regression and therefore more convenient to use here.
[14] This can actually be measured with a Grubel-Lloyd index for intra-industry trade. This index requires import and export data which is not available for all countries for all time periods included in this study and could therefore not be calculated.
[15] Note that these values are 'qualitative markers' rather than meaningful ordered quantities.
[16] page 782 of Wackerly et al (2008)

or transition and zero if not. The runs test is not done for ys itself, only for the separate phases. A Kolmogorov-Smirnov test[17] is performed to see whether the phases of the national cycles, ys, behave randomly. The same is done for the phases of the world cycle. The Kolmogorov-Smirnov test compares two samples under the null hypothesis that the two samples have the same distribution. The national and world cycles are compared against a randomly drawn sample which can take four values.

If the succession of business cycle phases is random according to the runs or Kolmogorov-Smirnov test and the null hypothesis of randomness cannot be rejected, modeling the business cycle phases and prediction may be nothing more than mere foolishness. If one or a few countries have a sequence that appears random whereas the others do not, then such countries may be marked as outliers and be left out of the analysis. Similarly, if the phase sequences appear to be non-random for the time period 1972-2004 but do appear so for the time period 2005-2008, then forecasting may be pointless regardless of the in-sample performance. Lastly, randomness in the cycle phases of the world business cycle will be examined. If this sequence appears to be random and is correlated to the national business cycles, then it may explain part of the randomness that might be found for the national sequences.

The correlation between the national business cycles and the world cycle is shown in table A1 in Appendix A, together with information per country about the national business cycles. The correlation seems roughly similar across countries, except for the negative correlation for New Zealand. As this is a correlation between phases, whose numbers are a qualitative marker rather than a meaningful quality, the correlations are complemented with a ‘hit-rate’. This hit-rate shows how often the phase of the world cycle and the phase of the national cycle are equal.

5. Methods

5.1 Length, Amplitude, Slope and Mass

Cycles can be seen as objects with their own characteristics, such as length, amplitude and mass. The interval unit changes as these characteristics exist per cycle rather than per year (note that the phases and transitions do exist per year). Explanatory variables for these dependent variables would be the averages during those cycles. The question in this case therefore slightly changes to: "How does the average self-employment during a cycle affect a cycle's characteristics?"

The number of cycles per country varies considerably which is why the panel structure is abandoned. The cycles are therefore deprived of a time and country context. For amplitude and mass OLS is used whereas for the length of the cycle Poisson models and NegBin models are used. As the latter were not part of the Bachelor's curriculum, the next section contains a short explanation. Some information about these cycles is shown in table A1 in Appendix A.

By definition, the average slope of a cycle is zero because the GDP data is de-trended. On a yearly basis the slope is already incorporated in the phases and therefore not explicitly incorporated as a dependent variable in a regression. The phases and transitions, are modeled with binary and multinomial logit models as explained in section 5.3.

[17] The Kolmogorov-Smirnov test is chosen over the χ^2 test based on the arguments put forward by Massey (1951). Alternatives could be the Wilcoxon rank sum test or the Mann-Whitney U-test, see Wackerly et al (2008).

5.2 Poisson and NegBin Models

The length of a cycle, measured in years, is always a positive integer larger than zero. It therefore fits the category 'count data'. The usual model would be a Poisson model.[18] Log transformations or Nonlinear Least Squares could be applied but are not ideal (see Wooldridge 2002). The Poisson regression assumes that the following three formulas hold.

$$\mu(x) \equiv E(y|x) \tag{7}$$

$$f(y|x) = e^{-\mu(x)} \frac{\mu(x)^y}{y!} \qquad y = 0, 1, \ldots \tag{8}$$

$$Var(y|x) = \sigma^2 E(y|x) \tag{9}$$

Formula (7) denotes the mean, (8) specifies the probability density function and (9) is an assumption about the variance. If the dependent variable truly follows a Poisson distribution, then σ^2 in (9) would equal one. In practice this assumption may be too restrictive. In (9) σ^2 is the mean-variance ratio and $\sigma^2 > 1$ indicates overdispersion, whereas $\sigma^2 < 1$ indicates underdispersion.

A common and popular parametric model $m(x, \beta)$ to estimate $\mu(x)$ is $m(x, \beta) = e^{-x\beta}$. The mean marginal effects (M.M.E.) can be estimated using (10).

$$\frac{\partial E(y|x)}{\partial x_j} = e^{\bar{x}\beta} \beta_j \tag{10}$$

The quasi-maximum likelihood estimator is robust to distributional misspecification. The formula for the robust standard errors is given by (12).

$$A_0 \equiv E[-H(\beta_0)] \tag{11}$$

$$Av\hat{a}r(\hat{\beta}) = \hat{\sigma}^2 \hat{A}^{-1} N^{-1} = \hat{\sigma}^2 (\textstyle\sum_{i=1}^{N} \nabla_\beta \widehat{m}_i' \nabla_\beta \widehat{m}_i / \widehat{m}_i)^{-1} \tag{12}$$

Where $\widehat{m}_i = m(x_i, \hat{\beta})$ and ∇_β is the gradient with regard to β. See Wooldridge (2002) for more details[19].

It could be that assumption (9) does not hold. Suppose that u_i is given by (13), then (14) would be an alternative model (see Wooldridge 2002 and 1991), where $h(x_i, \beta_0)$ is an alternative to $m(x_i, \beta_0)$.

$$u_i \equiv y_i - m(x_i, \beta_0) \tag{13}$$

$$E(u^2|x_i) = \sigma_0^2 m(x_i, \beta_0) + h(x_i, \beta_0)\delta_0 \tag{14}$$

Assumption (9) is rejected if the null hypothesis H_0: $\delta_0 = 0$ does not hold. If this is the case, NegBin models can replace the Poisson models. NegBin models use the same estimator for $\mu(x)$ and are similar in several

[18] This section is largely based on the explanation in Wooldridge (2002)
[19] The formulas for the robust standard errors are a rewritten form of the GMM sandwich estimator.

aspects except in that the NegBin models reckons with unobserved effects. Formulas 15, 16, 17 and 18 show the formulas pertaining to the NegBin models (Wooldridge 2002), where c_i denote the unobserved effects.

$$y_i \mid x_i , c_i \sim Poisson[c_i m(x_i , \beta)] \tag{15}$$

$$c_i \sim Gamma(a, b) \tag{16}$$

$$E(c_i) = 1 \ , \quad Var(c_i) = \eta^2 \tag{17}$$

$$\begin{aligned} Var(y_i|x_i) &= E[Var(y_i| x_i , c_i) \mid x_i] + Var[E(y_i| x_i , c_i)| x_i] \\ &= m(x_i, \beta) + \eta^2[m(x_i, \beta)]^2 \end{aligned} \tag{18}$$

Setting $h(x_i , \beta_0) = [m(x_i , \beta_0)]^2$ in (10) is convenient as then $\delta_0 = \eta^2$. This corresponds to the NegBin II model (Cameron and Trivedi 1986). Marginal effects and robust standard errors for a QMLE approach for the NegBin models are similar to those of the Poisson models.

Ordinarily, the negative binomial distribution is a generalization of the geometric distribution and if $Y \sim NegBin(p, r)$ then the probability density function at Y shows the probability of reaching r successes in Y trials if the probability of success is p. In the NegBin models the r is replaced by η^2 though there does not seem to be a meaningful interpretation of this. Rather it seems that utilizing the NegBin count data models is a 'technical issue' due to the variance specification.

5.3 Binary and Multinomial Logit Models

The business cycle phases and transitions are modeled using binary logit and multinomial logit models. In the binary logit models the phases and transitions are modeled separately and a ‘one’ would denote being in that phase or transition and a ‘zero’ otherwise. In the multinomial logit model the dependent variable can assume four values corresponding to the four phases. The choice between logit and probit models depends in part on whether there are random or fixed effects as explained earlier.
If there happens to be no heterogeneity, then the choice is quite trivial. In such a case the choice would be logit over probit because the probability and cumulative density functions have easier forms, allowing easier programming. There is a test to see whether the results between logit and probit models differ significantly, namely the Vuong test (Vuong 1989). Its formulas are shown in appendix B.

5.3.1 Binary Logit

The dependent variable is binary and modeled with the use of a latent variable y^* as in (19) and (20)[20]. The threshold c in (19) is commonly set as c = 0. Formula (21) holds for general functional forms $G(t)$ and (22) shows what the cumulative density function is if the functional form $G(t)$ is the logit function $\Lambda(t)$.

$$y = \begin{cases} 1 \ if \ y^* > c \\ 0 \ if \ y^* < c \end{cases} \tag{19}$$

[20] This section is largely based on the explanation in Heij et al (2004)

$$y^* = X\beta + \epsilon^* \tag{20}$$

$$P(Y = 1 \mid X) = G(X\beta) = p(X) \tag{21}$$

$$G(t) = \Lambda(t) = \int_{-\infty}^{t} \lambda(s)\, ds = \frac{e^t}{1+e^t} = \frac{1}{1+e^{-t}} \tag{22}$$

As the latent variable and the associated disturbance term ϵ^* are not observed there are no natural residuals. Artificial, pseudo, standardized residuals can be created with formula (23).

$$e^* = \frac{y-p}{\sqrt{p(1-p)}} \tag{23}$$

Model performance can be evaluated using so-called 'hit-rates'. Probabilities are defined as in table 1, where ŷ is the estimated value of y. The hit-rate is $h = p_{00} + p_{11}$. A test statistic Z based on the hit-rate as shown in (24) reveals whether the model performs better than random. In (24) it holds that $q = p^2 + (1-p)^2$.

	y = 0	y = 1	
ŷ = 0	p_{00}	p_{01}	$p_{0.}$
ŷ =1	p_{10}	p_{11}	$p_{1.}$
	$p_{.0}$	$p_{.1}$	1

Figure 1: general prediction-realization table

$$Z = \frac{h-q}{\sqrt{q(1-q)n}}, \ Z \sim N(0,1) \tag{24}$$

Here $\hat{y} = 1$ if $G(x_i'b) > c$, where c is usually taken to be 0.5 but could also be the fraction of ones in the sample, $c = \frac{\# y_i=1}{n}$. The number of observations is denoted by ‘n’.

Other measures to compare models would be McFadden’s pseudo R^2, Maddala’s R^2, McKelvey and Zavoina's R^2, the Akaike Information Criterion (AIC) and the Bayes Information Criterion (BIC) (Heij et al 2004, Maddala 1983, McKelvey and Zavoina 1975). Only McKelvey and Zavoina’s R^2 is shown, in (25), the others are shown in Appendix B. The McKelvey and Zavoina R^2 in (27) is used in this thesis because it is more convenient to calculate. In (25) $\bar{y}$ is the average of $\hat{y}$ and $\sigma^2 = \frac{1}{3}\pi^2$.

$$R^2 = \frac{\sum_{i=1}^{N}(\hat{y} - \bar{y})^2}{\sum_{i=1}^{N}(\hat{y} - \bar{y})^2 + N\sigma^2} \tag{25}$$

Once the best model has been chosen the coefficients need to be interpreted. There are few misspecification tests for the logit model. One could test for heteroskedasticity. The formulas are shown in Appendix B.

A way to improve the model and its interpretation is through robust variance matrices. A robust variance matrix for the coefficients can be calculated with (26) and (27). A robust variance matrix for the mean marginal effect δ_k as defined in (28) is given by (29) (see Wooldridge 2002).

$$A(X_i,\beta) \equiv -E[H_i(\beta)|X_i\,] = \frac{[g(X_i\beta)]^2 x_i' x_i}{G(X_i\beta)[1-G(X_i\beta)]} \tag{26}$$

$$\hat{V} \equiv Av\hat{a}r(\hat{\beta}) = \sum_{i=1}^{n} A(X_i,\beta) \tag{27}$$

$$\delta_k = \beta_k g(X\beta) = \frac{\partial P[Y=1|X]}{\partial x_k} \tag{28}$$

$$Var(\delta_k) = [\nabla_\beta \delta_k]\hat{V}[\nabla_\beta \delta_k]' \tag{29}$$

Similar methods and formulas exist for the multinomial logit model. Note that the mean marginal effects are evaluated at the average values of the explanatory variables, $\bar{x}$.

5.3.2 Multinomial Logit

The most important change when moving from binary to multinomial is transforming (21) into (30). The multinomial model has m categories[21]. Identification problems in the parameters β are solved by fixing $\beta_{1h} = 0, \forall\, h$. This ensures that the probabilities for all the categories sum to one. Another change is that in the hit-rate formulas: $\hat{q} = \sum_{j=1}^{m} \hat{p}_j^2$. and $h = \sum_{j=1}^{m} p_{jj}$.

$$P[y_i = j] = p_{ij} = \begin{cases} \dfrac{1}{1+\sum_{h=2}^{m} e^{X_i'\beta_h}} & voor\ j = 1 \\ \dfrac{e^{X\beta_j}}{1+\sum_{h=2}^{m} e^{X_i'\beta_h}} & voor\ j = 2, \dots, m \end{cases} \tag{30}$$

Formulas (31) and (32) are the odds ratios and the marginal effects respectively.

$$\frac{\partial p(x)/\,\partial x_j}{\partial p(x)/\,\partial x_h} = \frac{\beta_j}{\beta_h} \tag{31}$$

$$\frac{\partial P[Y_i = j\,|X_i]}{\partial x_i} = P[Y_i = j\,|X_i]\left(\beta_j - \sum_{h=2}^{m} p_{ih}\beta_h\right) \tag{32}$$

Odds ratios are more complicated in the multinomial case. It denotes the tendency of category j over h. This relies on the IIA assumption, the Independence of Irrelevance Alternatives. That means that a third category k does not influence the preference of category j over h. Whether this holds can be tested with a Hausman test as Franses and Paap (2004) demonstrate.

The log-likelihood changes as well, from (33) in the binary case to (34) in the multinomial case.

$$l(\beta) = \sum_{i=1}^{n}\sum_{t=1}^{T}\{y_{it}\log G(x_{it}\beta) + (1-y_{it})\log(1-G(x_{it}\beta))\} \tag{33}$$

$$l(\beta) = \sum_{i=1}^{n}\sum_{j=1}^{m} \mathrm{y_{ij}} \log p_{ij} \tag{34}$$

In (33) T is the number of time periods. The multinomial regressions do not have a panel structure. This can be seen when comparing formula (33) with (34) as (34) does not include a time dimension.

[21] This section is largely based on the explanation in Heij et al (2004)

5.4 Variable Selection

One of the main reasons to inspect the data with both a multinomial model and several separate binary models for each phase is variable selection. In the multinomial model the same explanatory variables are used for each phase, whereas the binary models allow entirely different sets of explanatory variables for each phase and transition. The binary models can be combined to forecast the state of the business cycle. The predicted state is then equal to the state j which has the highest predicted probability among all states, see (35).
In the multinomial model the probabilities of all states sum to one, which is not necessarily the case for combining the binary models.

$$\hat{y}_i = j \ if \ p_{ij} = \max_{\mathrm{h}}(p_{ih}) \quad (35)$$

The predictions will range from 2005 till 2008 and the time period 1972 to 2004 is used for calculating the coefficients in the models. The average cycle lasts for 5.26 years. Forecasting up to half a cycle ahead seems a reasonable upper limit, which means that for each variable three lags will be considered for inclusion for each binary model. As there are 81 variables, this would mean that for each binary model there are 243 candidates as explanatory variables and there are six such models, four for the phases and two for the transitions.
A common heuristic for variable selection is the general-to-specific method[22]: include all potential explanatory variables and drop the one with the highest P-value[23]. Repeat this with the remaining explanatory variables until all P-values are equal to or below the significance level[24]. As this is not practically feasible with 243 variables it happens in a five step procedure.

In the first step cycle characteristics are included. The variables that appear significant according to their P-values constitute the CYCH model (CYCH from 'cycle'). The second step builds on the CYCH model by adding[25] unemployment and self-employment variables. This creates the EMPL models (EMPL for 'employment'). The third step adds time and country dummies, interactions, nonlinear variants and produces the EXTE models (EXTE for 'extension'). The fourth step attempts additional macro-economic variables such as population, GDP per capita, average firm size in both workers and GDP and many more. This results in the MACR models (MACR for 'additional macroeconomic variables'). The fifth step incorporates world influence by adding to the CYCH, EMPL, EXTE and MACR models global, aggregate variables. This leads to the W.CYCH, W.EMPL, W.EXTE and W.MACR models (W. for 'world'). The global variables are added to all previous models to better understand the influence of the world given other variables, a central theme of this thesis.

5.5 Model Selection

The procedure in section 5.4 spawns 8 models for each phase and transition. Which model is best? This can be evaluated with the Akaike Information Criterion and the Bayes Information Criterion[26] which roughly speaking incorporates both improvement in R^2 and the number of variables employed. More explanatory variables lead to an increase in the R^2 but this does not mean that they should be added as there is the danger of overfitting[27].

---

[22] See Gilbert (1986) and Pagan (1987) for a discussion on such methods.
[23] The intercept is not dropped regardless of its P-value. Dropping the intercept would mean that the share of ones and zeros is equal. This is not the case here and Franses and Paap (2004, p.57) advise to keep the intercept.
[24] Significance level α = 0.05 unless stated otherwise
[25] Whether a group of additional variables as a whole has a significant influence, next to the explanatory variables already in the model, can be tested with an LR test. $LR = -2(l(\hat{\theta}_1) - l(\hat{\theta}_0)), \quad LR \sim \chi^2(g)$, where g is the number of additional variables and $l(\hat{\theta}_1)$ and $l(\hat{\theta}_0)$ are the log-likelihoods of the models with and without the additional variables respectively.
[26] Also known as the SIC, 'Schwarz Information Criterion'
[27] See Heij et al (2004) for more discussion on this topic.

One selection method would be to pick the model which has the lowest AIC or BIC or both. Generally the BIC tends to be more 'conservative' in that it points at models with as many or less explanatory variables than the AIC does. The formulas below are derived from Franses and Paap (2004) and are based on log-likelihood rather than R^2 or residuals as those are not be naturally defined in the case of binary and multinomial models.

$$AIC = \frac{1}{N}(-2l(\hat{\beta}) + 2g) \quad (36)$$

$$BIC = \frac{1}{N}(-2l(\hat{\beta}) + g \log N) \quad (37)$$

It should be noted that model selection is based on in-sample performance with the sample covering the time period 1972-2008.

5.6 Forecasting with a Moving Window and Sideward Replication

Once the model has been chosen it will be evaluated on its forecasting performance. Initially coefficients are estimated using the sample 1972-2004. This sample is the so-called 'window'. Those coefficients are used to forecast one year ahead, 2005. A two year ahead forecast is achieved by moving the window one year, which means that the coefficients of the model are estimated with the sample 1973-2005. Observations of the explanatory variables in 2005 are used in the forecast. Generally speaking, forecasting k years ahead means moving the entire 'window' k - 1 years ahead, which is why this method is called the 'moving window'.

Besides forecasting there will also be two sideward replications. This means that rather than dividing the sample in two across time it is divided in two across countries. In each replication, the phases of two countries are predicted based on the observations of the other twenty-one countries. In the first sideward replication those two countries are Iceland and Switzerland and in the second sideward replication those two are Finland and Spain[28]. It should be noted that the variable and model selection procedures as described in sections 5.4 and 5.5 use the information that is to be predicted. This only concerns the selection of which explanatory variables to use in the models though.

There is an implicit assumption not mentioned so far. The creation of forecasts using formula (35) does not explicitly incorporate dependency on previous phases. However, it is for example not possible to reach state BN from state RN. The model does not take this into account. The influence of previous phases will to some extent be dealt with by including for each phase lags of binary explanatory variables in the variable selection procedure. These variables then indicate whether previously the business cycle was in a certain phase. There is also another way to deal with the influence of previous states. This is shown in formula (36).

$$\hat{y}_i = j \ if \ p_{ij} = \max_{\mathrm{h}}(p_{ih} \cdot p^*_{h|\hat{y}_{i-1}}) \quad (36)$$

As in (35), p_{ih} denotes the probabilities as estimated by the binary or multinomial logit models. The difference is $p^*_{h|\hat{y}_{i-1}}$, which is the probability that the phase of the business cycle at time i will be phase h given that the phase at the previous year was $\hat{y}_{i-1}$. The probability that the phase at time i will be BN given that in the previous year it was RP will be zero as such a move is impossible. The probabilities $p^*_{h|\hat{y}_{i-1}}$ are estimated by examining how often transitions from one phase to another occur in a ten year moving window across countries. The forecast will be repeated with the method from formula (36) to see whether it will provide improved predictions.

[28] These countries were randomly selected with a pseudo random number generator.

6. Results

6.1 Runs test and Kolmogorov-Smirnov test

The runs test rejects the null hypothesis of randomness for all national binary variables, as a whole and for individual countries and in the whole time period and the forecast time period. The business world cycle tells a different story. When examining the whole time period, the test does not reject randomness for the binary variable indicating the phase RN, recession-negative slope. For the time period 2005-2008 this happens for the binary variables RP, RN and N (negative-slope)[29].
The Kolmogorov-Smirnov test rejects the null hypothesis that the phases of the national and world cycles have the same distribution as a randomly drawn sample where the observations can take four values. The test does not reject the null hypothesis that the phases of the national business cycles and the world business cycle have the same distribution.

6.2 OLS and Poisson Models

The total number of cycles in data is 121, with an average of 5.26 cycles per country. The standard deviation is 0.964, minimum number of cycles per country is 4 and the maximum 8. The results for the Poisson regressions and the Negative Binomial regressions are identical. In table A3 in Appendix A the cycle length is regressed on the amplitude during the cycle, the average self-employment rate and the average unemployment rate. In table 1 this regression is repeated with the addition of the global variables. ***M.M.E.*** is short for '*Mean Marginal Effect*'.

x	β	M.M.E.	σ, robust	P-value	Confidence interval	
constant	-0.724	-4.614	0.315	0.022	-1.342	-0.105
amplitude	4.826	30.759	1.517	0.001	1.853	7.8
average self-employment	-0.174	-1.109	0.371	0.64	-0.901	0.554
average unemployment	0.007	0.045	0.007	0.34	-0.007	0.021
global amplitude	34.407	219.293	3.355	0	27.831	40.983
global aver. self-employment	-7.323	-46.673	3.138	0.02	-13.473	-1.172
global aver. unemployment	41.642	265.405	2.909	0	35.939	47.344

Table 1: Poisson regression on the cycle length, global and national

These regressions show that the global variables have a significant influence on the length of the national cycles and national variables do not. The R^2 is 0.023 for the regression in table A3 and 0.242 in table 1.

Table A4 in Appendix A and the left part of table 2 below show regression on the mass of the cycles. The self-employment rate and unemployment rate do not attain significance on a national level but do on a global level. The R^2 remains low though with 0.026 and 0.082 for table A4 and 2 respectively. Repetition of these regressions for the cycle amplitude[30] yield table A5 in Appendix A and the right part of table 2. For these

[29] Although the phases 'boom', 'recession', 'positive slope' and 'negative slope' are not incorporated elsewhere, they were examined for randomness. The purpose of research vanishes if there is randomness, which is why it warrants more thorough examination.
[30] Expressed in percentages of total GDP

regressions none of the explanatory variables appear to have significant influence and R^2 is 0.02. This also holds if global variables are added.

x	β	σ, robust	P-value	β	σ, robust	P-value
constant	97806.16	76926.79	0.206	0.046	0.016	0.004
average self-employment	-161223.5	130203.3	0.218	0.017	0.027	0.521
average unemployment	2686.643	2593.219	0.302	0	0.001	0.36
global aver. self-employment	-1983722	834262.9	0.019	-0.27	0.172	0.119
global aver. unemployment	2585682	1094416	0.02	0.147	0.226	0.516

Table 2: OLS regression on the cycle mass (left) and amplitude (right), global and national variables

As far as misspecification is concerned, both regressions display similar results. The mean of the residuals in both regressions does not significantly differ from zero but the Jarque-Bera test rejects the null hypothesis that the residuals are normally distributed. The reason appears to be thick tails. Heteroskedasticity was not found (see Appendix B for a method to discover heteroskedasticity). Correlation between the explanatory variables and the residuals remains close to zero for both regressions and generally stay between -0.01 and 0.01. There is no indication for endogeneity.

The length and mass of national business cycles appears to be influenced by the self-employment rate and unemployment rate on a global level but not on a national level. This will be examined in more detail in the next section.

6.3 Binary Logit Models

It so happens that for all regressions the LR test rejected the presence of heterogeneity. Logit is therefore chosen over probit as explained earlier. The application of the five step procedure mentioned at section 5.4 yielded an interesting picture. The models vary widely in the number of explanatory variables as can be seen in table 3 and the categories of explanatory variables that are employed vary extraordinarily as well. The national self-employment rates were never significant, the global self-employment rates almost always were.

# variables	CYCH	EMPL	EXTE	MACR	W.CYCH	W.EMPL	W.EXTE	W.MACR
BP	7	9	10	13	26	28	23	24
BN	2	5	12	12	14	10	15	15
RP	4	5	10	10	7	7	11	12
RN	5	6	13	15	20	19	28	32
TBR	4	6	9	13	9	9	13	18
TRB	2	6	14	15	13	15	15	18

Table 3: number of variables in the logit models

Time dummies occasionally were significant though only the ones denoting decades, see also table A6 in Appendix A. Country dummies and dummies for groups of countries never managed to reach P-values below the significance level. There does not appear to be an indication that the results hinge crucially on one country or a small group of countries.

The variation across models is reflected in the (pseudo-) R^2 as shown in table 4 which increase with models that cover more potential explanatory variables. This is McKelvey and Zavoina's R^2. Addition of the global variables enhances the R^2 considerably. Yet this comes at the cost of more explanatory variables. Whether the increase in R^2 justifies the increase in explanatory variables is revealed by the AIC and BIC in tables A9 and A10 in Appendix A.

R^2	CYCH	EMPL	EXTE	MACR	W.CYCH	W.EMPL	W.EXTE	W.MACR
BP	0.392	0.395	0.534	0.581	0.583	0.596	0.666	0.705
BN	0.565	0.665	0.909	0.909	0.766	0.762	0.9	0.9
RP	0.661	0.678	0.905	0.914	0.736	0.736	0.922	0.918
RN	0.149	0.158	0.265	0.285	0.363	0.352	0.443	0.458
TBR	0.278	0.283	0.868	0.882	0.411	0.411	0.859	0.881
TRB	0.122	0.167	0.699	0.852	0.444	0.44	0.747	0.888

Table 4: McKelvey and Zavoina's R^2 for the logit models

These tables show that the AIC selects W.MACR as the best of the available models 5 out of 6 times and that the more conservative BIC sometimes chooses W.EXTE and sometimes W.MACR. This happens despite the fact that those models employ the largest numbers of explanatory variables. As table A7 in the appendix shows, all models achieve respectable hit-rates and all of them are better than random according to the test in (24). The W.MACR models will therefore be used as the models of choice for forecasts and sideward replication. The Vuong test did not reject the null hypothesis that the results for the logit models differ from the results of a similar probit model.

6.4 Multinomial Logit Models

The multinomial logit models look at all four phases at once and the explanatory variables are the same for all phases. Their P-values vary considerably per phase so variables will not be selected on those P-values. Rather, the multinomial model uses the first three lags of national cyclical GDP, the national self-employment rate, the national unemployment rate, the phases from the world cycle and the global self-employment rate.

The R^2 of the regression is 0.364, which is quite low compared to those of the binary logit variables. The results of regression, coefficients and their P-values, are given in table A11 in Appendix A. Note that the table does not contain results for the phase BP as this is taken as the base category that is set to zero to avoid identification problems. The multinomial logit model does seem to be inferior to the binary logit models as many variables are not significant in multiple phases. The only misspecification test that applies to multinomial logit would be the IIA assumption but this assumption cannot hold as it is impossible to reach every phase from a given phase.

6.5 Forecasts

The forecast results are shown below in table 5. The initial estimation sample is the time period 1972-2004 and the time period 2005-2008 is forecasted. The separate binary logit models seem to predict better than random in four out of six cases according to the test in (24). Striking is the result for W.MACR-BP. It occurred substantially more often in the forecast sample than in the estimation sample, 48.9% of the time against 30.9% of the time. This could be due to the build-up of the real estate bubble. The tables in A12 in Appendix A show the prediction-realization tables for each of the six forecast models separately.

Model	Hit-rate better than random?	Heteroskedasticity found?	C = (#y=1 / n) in estimation sample	C = (#y=1 / n) in forecast sample
W.MACR -BP	No, it is 0.544	No	0.309	0.489
W.MACR -BN	Yes, it is 0.848	No	0.164	0.13
W.MACR –RN	No, it is 0.598	No	0.313	0.272
W.MACR -RP	Yes, it is 0.935	No	0.215	0.109
W.MACR -OBR	Yes, it is 0.837	No	0.132	0.163
W.MACR -ORB	Yes, it is 0.87	No	0.119	0.174

Table 5: forecast results binary logit models

Combination of the binary logit models for the four phases to create phase predictions as explained at 5.4 yields the left part of prediction-realization table 6 below. The hit-rate is 0.446 and it is not significantly better than a random draw according to the generalized hit-rate test. This is probably due to the fact that the binary logit model struggles to predict the Boom-Positive slope (BP) phase.

	y=BP	y=BN	y=RN	y=RP	Total	y=BP	y=BN	y=RN	y=RP	Total
ŷ = BP	0.293	0.109	0.141	0.011	0.554	0.283	0.087	0.12	0.011	0.5
ŷ = BN	0.011	0.022	0.033	0	0.065	0.033	0.033	0.043	0.011	0.12
ŷ = RN	0.163	0	0.087	0.054	0.304	0.141	0.011	0.087	0.011	0.25
ŷ = RP	0.022	0	0.011	0.043	0.076	0.033	0	0.022	0.076	0.13
Total	0.489	0.130	0.272	0.109	1	0.489	0.13	0.272	0.109	1

Table 6: prediction-realization table phase predictions, binary logit models (left) and multinomial logit model (right)

Interestingly, the multinomial model performs equally well as the binary logit models and even has a slightly higher hit-rate of 0.478. The results of the multinomial model are shown in the right part of table 6. The multinomial model struggles with the BP phase as well and predicts it about equally well as the binary logit models. The binary logit models were built up carefully using a wide range of 243 possible explanatory variables and had high R^2 and hit-rates, whereas the multinomial model has a few explanatory variables which are often not significant and has a relatively low R^2. This suggests that there are deeper structures and dynamics not captured by the employed models.

Table 7 shows the results when an adjustment has been made to incorporate the previous phase of the cycle, as explained with formula (36) in section 5.6. Again the null hypothesis that the models forecast equally well as a random draw is not rejected and the hit-rates are 0.457 for the binary models and 0.424 for the multinomial model. The models have become more adapt at forecasting the phase BP but worse at forecasting the other phases. Lags of phases were not significant as explanatory variables so this result does not come as a surprise.

	y=BP	y=BN	y=RN	y=RP	Total	y=BP	y=BN	y=RN	y=RP	Total
ŷ = BP	0.38	0.098	0.228	0.043	0.75	0.326	0.098	0.174	0.033	0.63
ŷ = BN	0.011	0.022	0	0	0.033	0.054	0.011	0.022	0.022	0.109
ŷ = RN	0.065	0.011	0.011	0.022	0.109	0.098	0.022	0.043	0.011	0.174
ŷ = RP	0.033	0	0.033	0.043	0.109	0.011	0	0.033	0.043	0.087
Total	0.489	0.13	0.272	0.109	1	0.489	0.13	0.272	0.109	1

Table 7: predict.-realiz. table adjusted phase predictions, binary logit models (left) and multinomial logit model (right)

6.6 Sideward Replication

Two sideward replications were carried out. Sideward replication 1 involved predicting the phases of Iceland and Switzerland and tables 8 and 9 contain the results, while sideward replication 2 involved Finland and Spain and its results are displayed in tables A13 and A14 in Appendix A.

Model	Hit-rate better than random?	Heteroskedasticity found?	C = (#y=1 / n) in estimation sample	C = (#y=1 / n) in forecast sample
W.MACR -BP	Yes, it is 0.868	No	0.332	0.309
W.MACR -BN	Yes, it is 0.883	No	0.16	0.162
W.MACR –RN	No, it is 0.635	No	0.308	0.309
W.MACR -RP	Yes, it is 0.838	No	0.2	0.221
W.MACR -OBR	No, it is 0.75	No	0.136	0.132
W.MACR -ORB	No, it is 0.794	No	0.125	0.132

Table 8: sideward replication 1 results for the binary logit models

	y=BP	y=BN	y=RN	y=RP	Total	y=BP	y=BN	y=RN	y=RP	Total
ŷ = BP	0.191	0	0.029	0	0.221	0.191	0	0.088	0.015	0.294
ŷ = BN	0.015	0.103	0.029	0	0.147	0.044	0.132	0.015	0	0.191
ŷ = RN	0.059	0.059	0.206	0	0.324	0.015	0.029	0.162	0	0.206
ŷ = RP	0.044	0	0.044	0.221	0.309	0.059	0.000	0.044	0.206	0.309
Total	0.309	0.162	0.309	0.221	1	0.309	0.162	0.309	0.221	1

Table 9: predict.-realizat. table sideward replication 1, binary logit models (left) and multinomial logit model (right)

The binary and multinomial logit models do not perform better than random, despite hit-rates of 0.721 for the logit model and 0.691 and 0.662 for the multinomial model. This is striking as for sideward replication 2 the binary logit models succeed in predicting all the phases separately better than random. Again as in the previous section it is striking how close the hit-rates of the binary and multinomial logit models are. This shows that a multinomial model based on entrepreneurship and with many insignificant variables does not have to yield substantially inferior predictions than binary logit models with carefully selected, significant variables.

7. Discussion and Conclusion

The previous sections affirm the findings from Koellinger and Thurik (2012) that the national business cycles are not significantly influenced by the national self-employment rates. The main research question of this thesis was:

“How do national business cycle characteristics depend on entrepreneurship?”

Generally, the global self-employment rate does seem to influence national business cycles, whereas the national self-employment rates do not. This is the case for the length of the cycle, for ‘mass’ and for the phases and transitions. However, the amplitude itself was not significantly influenced by national or global self-employment rates. The national self-employment rates never were significant in the selection of variables at any stage whereas the global self-employment rate was.

The forecasts and sideward replications revealed a surprising picture. The global and national self-employment rates predicted about equally well as carefully crafted models with many other, more significant macroeconomic variables. This happened despite apparent inferior performance of global and national self-employment in-sample. A sizeable number of macroeconomic variables besides the self-employment rate were not able to create predictions better than random.

There do seem to be fundamental differences between the separate business cycle phases but the forecasts and sideward replications suggests that these may not be relevant for prediction of business cycle as a whole. Remarkable is the fact that despite the vast array of explanatory variables the phases and transitions differ substantially in the extent to which they can be explained and forecasted. The influence of the explanatory variables varies considerably and change per phase. Yet the inability to forecast better than randomly casts doubt on whether these differences are important. This could be caused by unobserved trends such as the real estate bubble. The boom-positive slope phase showed aberrant behavior during the forecast time period as opposed to the estimation sample time period, which probably wrecked the predictive qualities of the models.

The objective was to recommend to a politician whether or not to stimulate entrepreneurship to influence the national business cycle. The results in this paper suggest that stimulating the national self-employment rate may not result in different behavior in the national business cycle. The global self-employment rate does affect the national business cycle if only by increasing their length and ‘mass’.
A European politician who wants cycle lengths to be longer may therefore find it worth his while to try to stimulate entrepreneurship on a European level rather than a country level. As the ‘world business cycle’ was to some extent a proxy to incorporate trading relationships across the globe there is also another economic interpretation. This would be that politicians should focus on strengthening trading relationships with countries which encourage entrepreneurship and have increasing self-employment rates.
However, this holds for the business cycle length. Amplitude was not significantly influenced by self-employment rates, neither on the national or global level. This thesis does not offer an indication that entrepreneurship leads to ‘bigger booms’ or less severe recessions.

To return to the title of this thesis: the optimistic view of the entrepreneur does not *predict* the phase of the business cycle but the optimistic view of the global entrepreneur does *influence* the phase of the national business cycle.

Appendix A: Additional Tables

Country	Phase correlation with phases world cycle	'hit-rate': fraction of national phase equals world phase	Random according to Runs test?	Number of cycles	Average length of cycles	Standard deviation length of cycles
Austria	0.422	0.514	No	13	2.846	1.676
Belgium	0.711	0.73	No	12	3.083	1.929
Denmark	0.278	0.459	No	9	4.111	1.691
Finland	0.461	0.486	No	9	4.111	1.764
France	0.779	0.676	No	12	3.083	1.73
Germany	0.492	0.568	No	9	4.111	1.965
Greece	0.306	0.486	No	11	3.364	2.618
Ireland	0.309	0.459	No	8	4.625	2.973
Italy	0.614	0.622	No	12	3.083	1.975
Luxembourg	0.64	0.568	No	9	4.111	2.028
The Netherlands	0.54	0.568	No	10	3.7	2.111
Portugal	0.677	0.622	No	8	4.625	1.408
Spain	0.481	0.486	No	8	4.625	2.56
Sweden	0.322	0.541	No	10	3.7	2.003
United Kingdom	0.538	0.568	No	8	4.625	3.204
Iceland	0.571	0.622	No	9	4.111	1.27
Norway	0.122	0.405	No	9	4.111	2.315
Switzerland	0.552	0.622	No	11	3.364	1.748
USA	0.411	0.622	No	9	4.111	1.9
Japan	0.256	0.486	No	9	4.111	2.571
Canada	0.35	0.568	No	10	3.7	2.497
Australia	0.417	0.432	No	15	2.467	1.959
New Zealand	-0.441	0.108	No	8	4.625	1.996
World	1	1	Sometimes	5	7.4	3.578

Table A1: Cycle data per country

Country	# BP	# BN	# RN	# RP	Country	# BP	# BN	# RN	# RP
Austria	13	4	12	8	Spain	13	6	12	6
Belgium	13	5	12	7	Sweden	14	5	11	7
Denmark	14	5	12	6	United Kingdom	13	6	8	10
Finland	13	6	10	8	Iceland	13	9	7	8
France	13	5	10	9	Norway	12	6	12	7
Germany	11	9	11	6	Switzerland	12	4	14	7
Greece	12	6	12	7	USA	17	4	10	6
Ireland	13	7	10	7	Japan	13	2	16	6
Italy	11	5	12	9	Canada	14	8	10	5
Luxembourg	12	8	11	6	Australia	16	2	11	8
The Netherlands	12	7	10	8	New Zealand	14	7	10	6
Portugal	12	7	9	9	World	12	6	10	9

Table A2: cycle phase data per country

x	β	Mean Marginal Effects	σ, robust	P-value	Confidence interval	
constant	1.412	0.001	0.136	0	1.145	1.678
amplitude	15.126	0.028	2.157	0	10.899	19.353
average self-employment	-1.091	0.298	0.683	0.11	-2.429	0.247
average unemployment	0.048	-0.022	0.01	0	0.027	0.068

Table A3: Poisson regression on the cycle length, national variables

x	β	σ, robust	z	P-value	Confidence interval	
constant	28955.35	22656.18	1.28	0.204	-15910.06	73820.76
average self-employment	-149751.8	128895.4	-1.16	0.248	-404999.8	105496.3
average unemployment	3666.915	2399.036	1.53	0.129	-1083.829	8417.66

Table A4: OLS regression on the cycle mass, national variables

x	β	σ, robust	z	P-value	Confidence interval	
constant	0.025	0.005	5.52	0	0.016	0.034
average self-employment	0.025	0.026	0.96	0.34	-0.027	0.077
average unemployment	-0.001	0	-1.26	0.21	-0.002	0

Table A5: Regression on the cycle amplitude, national variables

Explanatory variables in the best logit models (W.MACR).

BP	Cycle slope lag 1 Boom-positive lag 1 70s dummy GDP cycle % squared lag 1 Unemployment rate squared lag 1 en 2 GDP per capita lag 2 and 3 GDP per capita cycle % lag 2 World GDP per capita lag 1 World GDP cycle % lag 2 World mass lag 1 and 2 World unemployment rate lag 1 and 2 World self-employment rate lag 1, 2 and 3 World unemployment rate cycle lag 1 and 2 World self-employment rate cycle lag 1, 2 and 3
BN	GDP Cycle % lag 1 Unemployment rate cycle lag 2 and 3 Boom-negative lag 2 Interaction GDP cycle % and self-employment rate cycle lag 1 and 2 80s dummy GDP Cycle % squared lag 1 World GDP cycle % lag 2 World slope lag 3 World unemployment rate lag 1, 2 and 3 World self-employment rate lag 1 World unemployment rate cycle lag 1 and 3
RP	GDP Cycle % lag 1 Cycle slope lag 1 GDP cycle % squared lag 1 and 3 Boom negative lag 3 World GDP cycle % lag 2 and 3 World slope lag 2 World mass lag 3 World unemployment rate lag 3 World self-employment rate lag 3
RN	GDP Cycle % lag 2 Unemployment rate lag 2 Recession-negative lag 3 Boom-negative lag 3 Boom-positive lag 1 70s dummy GDP cycle % squared lag 1 Unemployment rate squared lag 1 Unemployment rate cycle squared lag 2 Labor productivity lag 2 and 3 Total labor force lag 2 and 3 World GDP per capita lag 1, 2 and 3 World GDP cycle % lag 2 World slope lag 2 and 3 World mass lag 2 and 3 World unemployment rate lag 1, 2 and 3 World self-employment rate lag 1, 2 and 3 World unemployment rate cycle lag 1 and 2 World self-employment rate cycle lag 1 and 2
TBR	GDP Cycle % lag 1 en 2 Unemployment rate cycle lag 3

	GDP cycle % squared lag 2 Unemployment rate squared lag 1 and 3 Unemployment rate cycle squared lag 2 Labor productivity lag 2 and 3 Total labor force lag 2 and 3 World GDP per capita lag 2 World unemployment rate lag 1 and 2 World unemployment rate cycle lag 1 and 2 World self-employment rate cycle lag 2
TRB	GDP Cycle % lag 3 Cycle slope lag 2 Transition recession boom lag 2 Transition boom recession lag 1 en 2 Interaction GDP cycle % and self-employment rate cycle lag 3 GDP cycle % squared lag 1,2 and 3 GDP per capita cycle % lag 1 and 3 World GDP cycle % lag 2 and 3 World slope lag 2 World unemployment rate lag 2 and 3 World self-employment rate cycle lag 2

Table A6: explanatory variables used in W.MACR models

Hit-rates	CYCH	EMPL	EXTE	MACR	W.CYCH	W.EMPL	W.EXTE	W.MACR
BP	0.721	0.725	0.772	0.763	0.781	0.786	0.816	0.817
BN	0.871	0.875	0.894	0.894	0.9	0.893	0.904	0.904
RP	0.84	0.853	0.87	0.861	0.857	0.857	0.876	0.868
RN	0.71	0.715	0.765	0.763	0.772	0.765	0.803	0.794
TBR	0.853	0.858	0.873	0.876	0.859	0.859	0.889	0.893
TRB	0.872	0.871	0.885	0.88	0.88	0.871	0.885	0.895

Table A7: hit-rates for the logit models

Odd ratios	CYCH	EMPL	EXTE	MACR	W.CYCH	W.EMPL	W.EXTE	W.MACR
BP	0.493	0.492	0.492	0.493	0.492	0.493	0.492	0.493
BN	0.177	0.19	0.19	0.19	0.19	0.19	0.19	0.19
RP	0.246	0.246	0.253	0.29	0.248	0.248	0.251	0.253
RN	0.446	0.446	0.446	0.446	0.446	0.446	0.446	0.445
TBR	0.159	0.157	0.157	0.157	0.15	0.15	0.157	0.157
TRB	0.139	0.143	0.143	0.143	0.143	0.143	0.143	0.143

Table A8: odd ratios for the logit models at average values of the explanatory variables

AIC	CYCH	EMPL	EXTE	MACR	W.CYCH	W.EMPL	W.EXTE	W.MACR
BP	873.735	870.960	778.981	774.003	746.739	743.241	703.371	***694.212***
BN	484.380	422.461	382.011	382.011	391.455	388.249	362.411	***362.411***
RP	488.711	484.090	432.596	432.596	462.972	462.972	***417.291***	421.895
RN	891.994	890.249	842.294	841.973	821.399	823.162	782.170	***774.801***
TBR	557.181	536.203	459.929	453.506	526.565	526.565	432.808	***427.646***
TRB	566.128	555.182	467.046	449.731	518.708	513.039	438.364	***412.846***

Table A9: AIC for the logit models

BIC	CYCH	EMPL	EXTE	MACR	W.CYCH	W.EMPL	W.EXTE	W.MACR
BP	911.03	917.579	830.261	839.269	872.609	878.435	815.256	***810.759***
BN	498.537	450.432	442.615	442.615	461.383	439.53	***437.001***	***437.001***
RP	512.165	512.235	483.876	483.876	500.267	500.267	***473.234***	482.499
RN	915.303	922.882	***902.898***	916.563	919.298	916.4	917.364	928.642
TBR	580.636	568.836	506.547	518.772	573.473	573.473	***498.074***	516.221
TRB	580.2	587.815	536.974	524.32	583.974	587.629	512.954	***501.421***

Table A10: BIC for the logit models

Explanatory variable	β BN	P-value BN	β RN	P-value RN	β RP	P-value RP
Constant	-6.487	0	-2.937	0.02	-7.341	0
Cyclical component GDP; 1 Lag	92.717	0	-52.093	0	-140.823	0
Cyclical component GDP; 2 Lags	-2.342	0.892	38.156	0.006	23.127	0.13
Cyclical component GDP; 3 Lags	9.937	0.49	8.043	0.324	6.586	0.53
Self-employment rate; 1 Lag	-5.927	0.836	-2.971	0.908	-17.519	0.59
Self-employment rate; 2 Lags	32.552	0.499	27.927	0.514	23.364	0.642
Self-employment rate; 3 Lags	-29.636	0.421	-24.858	0.403	-6.762	0.834
Unemployment rate; 1 Lag	0.175	0.498	0.398	0.055	0.187	0.433
Unemployment rate; 2 Lags	-0.132	0.723	-0.222	0.484	-0.015	0.97
Unemployment rate; 3 Lags	-0.009	0.971	-0.164	0.367	-0.116	0.631
Global self-employ-ment rate; 1 Lag	25.684	0.671	-19.134	0.712	89.034	0.299
Global self-employ-ment rate; 2 Lags	38.936	0.552	45.139	0.424	-209.608	0.077
Global self-employ-ment rate; 3 Lags	-4.851	0.93	16.526	0.718	168.242	0.051
Phase World Business Cycle; 1 Lag	0.214	0.242	0.189	0.075	0.382	0.016
Phase World Business Cycle; 2 Lags	-0.553	0.003	-0.489	0	-0.152	0.421
Phase World Business Cycle; 3 Lags	-0.372	0.005	-0.394	0	-0.143	0.377

Table A11: coefficients and P-values for the multinomial model

W.MACR – BP	y = 0	y = 1	Total
ŷ = 0	0.283	0.228	0.511
ŷ = 1	0.228	0.261	0.489
Total	0.511	0.489	1

W.MACR – BN	y = 0	y = 1	Total
ŷ = 0	0.804	0.087	0.891
ŷ = 1	0.065	0.043	0.109
Total	0.87	0.13	1

W.MACR – RN	y = 0	y = 1	Total
ŷ = 0	0.522	0.196	0.717
ŷ = 1	0.207	0.076	0.283
Total	0.728	0.272	1

W.MACR – RP	y = 0	y = 1	Total
ŷ = 0	0.87	0.043	0.913
ŷ = 1	0.022	0.065	0.087
Total	0.891	0.109	1

W.MACR – TBR	y = 0	y = 1	Total
ŷ = 0	0.837	0.163	1
ŷ = 1	0	0	0
Total	0.837	0.163	1

W.MACR – TRB	y = 0	y = 1	Total
ŷ = 0	0.826	0.13	0.957
ŷ = 1	0	0.043	0.043
Total	0.826	0.174	1

Table A12: prediction-realization tables for W.MACR forecasts

Model	Hit-rate better than random?	Heteroskedasticity found?	C = (#y=1 / n) in sample	C = (#y=1 / n) in forecast sample
W.MACR -BP	Yes, it is 0.853	No	0.331	0.324
W.MACR -BN	Yes, it is 0.853	No	0.160	0.162
W.MACR –RN	Yes, it is 0.779	No	0.307	0.324
W.MACR -RP	Yes, it is 0.824	No	0.2	0.191
W.MACR -TBR	No, it is 0.824	No	0.137	0.118
W.MACR -TRB	No, it is 0.838	No	0.128	0.103

Table A13: sideward replication 2 results for the binary logit models

	y=1	y=2	y=3	y=4	Total	y=1	y=2	y=3	y=4	Total
ŷ = 1	0.265	0.015	0.044	0.029	0.353	0.235	0.029	0.029	0.029	0.324
ŷ = 2	0.044	0.132	0.015	0	0.191	0.029	0.118	0.015	0	0.162
ŷ = 3	0.015	0.015	0.176	0.015	0.221	0.044	0.015	0.176	0.029	0.265
ŷ = 4	0	0	0.088	0.147	0.235	0.015	0	0.103	0.132	0.25
Total	0.324	0.162	0.324	0.191	1	0.324	0.162	0.324	0.191	1

Table A14: sideward replication 2, binary logit models (left) and multinomial logit model(right)

Appendix B: Additional Formulas and Theory

Choosing a value for λ in the HP-filter

Ravn and Uhlig (2002) show with Fourier transformations that a value of λ = 6.25 is more appropriate than the standard value λ =100 for yearly.

Koellinger and Thurik (2012) showed that their results were not sensitive to the method of de-trending. They show their results with λ = 100, λ = 6.25 and with 'first differences'. The standard value for λ for yearly data in the literature remains 100. Visual inspection of the trend or growth component for all countries suggests that with the data employed here, a value of λ = 100 is preferable over λ =6.25. As can be seen in Figures B.1 and B.2, when λ = 6.25 the trend may display cyclical tendencies. The blue line is the 'raw' GDP which consists of both the trend and the cyclical component and the red line is the trend as created by the HP-filter.

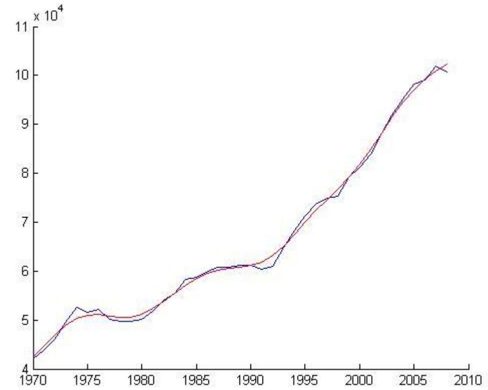

Figure B.1: GDP New Zealand (λ =6.25)

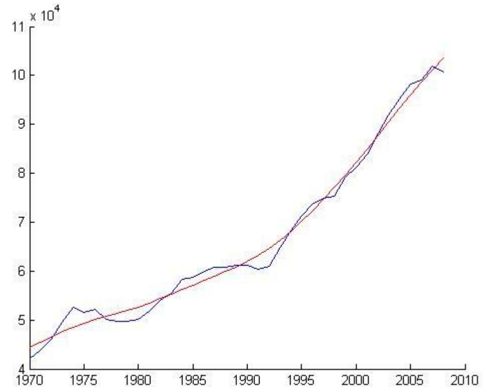

Figure B.2: GDP New Zealand (λ = 100)

The above considerations have led to the choice of using λ = 100 to de-trend the data in the HP-filter for this thesis.

A reason why the value of λ =6.25 may not be appropriate could be one of the underlying assumptions. Hodrick and Prescott (1980 and 1997) derive their value of λ for quarterly data on the argument that a five per cent deviation from the trend and an eight per cent change in the trend component are moderate. Ravn and Uhlig (1997) address this point as well, remarking that it is mainly about defining what one regards as the appropriate length of the business cycles. This seem to imply that it is also about defining what an appropriate trend component would be as that is the other side of the coin. The trend component in Figure B.2 seems more appropriate than the one in Figure B.1 but this is an opinion.

Additional formulas for logit models

In the formulas below, (B.1) equals the probability density function of Y. (B.2) is the derivative of $G(t)$ of earlier formulas, (B.3) shows McFadden's pseudo R^2 and (B.4) shows Maddala's pseudo R^2.

$$f(y\,|x_i,\beta) = [G(x_i\beta)]^y[1 - G(x_i\beta)]^{1-y} \qquad y \in \mathcal{B} \qquad \text{(B.1)}$$

$$g(t) = \lambda(t) = \frac{e^t}{(1+e^t)^2} \quad \text{(B.2)}$$

$$R^2 = 1 - \frac{l(\hat{\theta})}{l(\hat{\theta}_0)} \quad \text{(B.3)}$$

$$R^2 = \frac{1-\left(\frac{l(\hat{\theta})}{l(\hat{\theta}_0)}\right)^{\frac{2}{n}}}{1-\left(l(\hat{\theta}_0)\right)^{\frac{2}{n}}} \quad \text{(B.4)}$$

In these formulas $l(\hat{\theta})$ is the log-likelihood of the model and $l(\hat{\theta}_0)$ the log-likelihood of the restricted model with only the intercept.
The odds-ratio in (B.5) reveals the tendency of y = 1 over y = 0, marginal effects in (B.6) disclose the effect of a particular explanatory variable on the dependent variable and elasticity (B.7) unveil the percentage change in the dependent variable if a particular independent variable changes a per cent.

$$\frac{P[Y_i=1\,|X_i]}{P[Y_i=0\,|X_i]} = \frac{G(x_i'\beta)}{1-G(x_i'\beta)} \quad \text{(B.5)}$$

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial P[y_i=1]}{\partial x_{ji}} = \beta_j \frac{1}{n}\sum_{i=1}^{n} G(x_i'\beta)\,(1-G(x_i'\beta)) \quad \text{(B.6)}$$

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial P[Y_i=1\mid X_i]}{\partial x_{ji}} = \frac{1}{n}\sum_{i=1}^{n} P(Y_i=1\mid X_i)(1-P(Y_i=1\mid X_i))\beta_k\, x_{ji} \quad \text{(B.7)}$$

Testing for heteroskedasticity

Suppose that the standard deviation is of the form in (B.8). The auxiliary regression in (B.9) allows calculation of R^2_{nc} in (B.10), which is the non-centered R^2. Here z_i is a vector of variables without the constant term, $\hat{p}_i = G(x_i'b)$ and $\hat{e}_i^*$ is the residual from regression (B.9).

$$\sigma_i = e^{z_i'\gamma} \quad \text{(B.8)}$$

$$e_i^* = \frac{g(x_i'b)}{\sqrt{\hat{p}_i(1-\hat{p}_i)}}x_i'\delta_1 + \frac{g(x_i'b)x_i'b}{\sqrt{\hat{p}_i(1-\hat{p}_i)}}z_i'\delta_2 + \eta_i \quad \text{(B.9)}$$

$$R^2_{nc} = \frac{\sum_{i=1}^{n}(\hat{e}_i^*)^2}{\sum_{i=1}^{n}(e_i^*)^2} \quad \text{(B.10)}$$

Under the null hypothesis of homoskedasticity $H_0\!:\gamma = 0$, implying that equation (B.11) holds, where g is the number of parameters in γ.

$$LM = nR^2_{nc} \sim \chi^2(g) \quad \text{(B.11)}$$

The choice for z_i is crucial in this test.

Vuong test

The Vuong test examines whether the results from the logit model differ significantly from results of the probit model (Vuong 1989). The test statistic is given in formula (B.12).

$$T = \frac{l_{logit}(\hat{\beta}) - l_{probit}(\hat{\beta})}{\hat{\omega}_n \sqrt{n}}, \quad T \sim N(0,1) \tag{B.12}$$

$$\hat{\omega}_n = \frac{1}{n}\sum_{i=1}^{n} \ln\left(\frac{f_{logit}(y_i|\hat{\beta})}{f_{probit}(y_i|\hat{\beta})}\right)^2 - \left(\frac{1}{n}\sum_{i=1}^{n} \ln\left(\frac{f_{logit}(y_i|\hat{\beta})}{f_{probit}(y_i|\hat{\beta})}\right)\right)^2 \tag{B.13}$$

In these formulas, $l(\hat{\beta})$ denotes the maximized log-likelihood and $f(y_i|\hat{\beta})$ denotes the density function of y_i at $\hat{\beta}$.

References

Ahumada, H. and Garegnani, M. L. (2000). “*Assessing HP-filter performance for Argentina and the U.S. Macro Aggregates*,” Journal of Applied Economics, 3(2), 257-284

Andrews, W. K. D. (1988). "*Chi-Square Tests For Econometric Models: Introduction and Applications*," Journal of Econometrics, 37, 135-156.

Audretsch, D.B., Grilo, I. and Thurik, A. R.(2007). "Handbook of Research on Entrepreneurship Policy," Edward Elgar Publishing Limited: Cheltenham.

Bernanke, B. and Gertler, M. (1989). "A*ssociation Agency Costs, Net Worth, and Business Fluctuations*," The American Economic Review, 79(1), 14-31.

Birchenhall, C.R., Jessen, H., Osborn, D. R. and Simpson, P. (1999)."*Predicting U.S. Business-Cycle Regimes*," Journal of Business and Economic Statistics, 17(3), 313-323

Burda, M. and Wyplosz, C. (2005). “Macroeconomics: A European Text,” Oxford University Press: Oxford.

Caballero, R. J. and Hammour, M. L. (1991). "*The Cleansing Effect of Recessions*," NBER Working Paper No. 3922.

Camerer, C. and Lovallo, D. (1999). “*Overconfidence and Excess Entry: An Experimental Approach*,” The American Economic Review, 89(1), 306-318.

Cameron, A. C. and Trivedi, P. K. (1986). "*Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators and Tests*," Journal of Applied Econometrics, 1(1), 29-53.

Carlstrom, C. and Fuerst, T. (1997). “*Agency costs, net worth, and business fluctuations: a computable general equilibrium analysis*,” American Economic Review, 87, 893-910.

Chamberlain, G. (1980), ‘‘*Analysis of Covariance with Qualitative Data*,’’ Review of Economic Studies 47, 225–238.

Congregado, E., Golpe, A. A. and Parker, S. (2009). "*The Dynamics of Entrepreneurship: Hysteresis, Business Cycles and Government Policy*," IZA Discussion Paper No. 4093.

Corriveau, L. (1994). "*Entrepreneurs, Growth and Cycles*," Economica, New Series, 61(241), 1-15.

Eichenbaum, M. (1991). "*Real business-cycle theory: wisdom or whimsy?*" Journal of Economic Dynamics and Control, 15, 607-626.

Fairly, R. (2011). "Entrepreneurship, Economic Conditions, and the Great Recession," Discussion Paper No. 5725.

Faria, J. R., Cuestas, J. C. and Gil-Alana, L. A. (2009). "*Unemployment and entrepreneurship: A cyclical relation?*" Economics Letters, 105, 318–320.

Filardo, A.J. (1994). "*Business-Cycle Phases and Their Transitional Dynamics*," Journal of Business and Economic Statistics, 12(3), 299-308.

Forni, M. and Reichlin, L. (1998). "*Let's Get Real: A Factor Analytical Approach to Disaggregated Business Cycle Dynamics*," Review of Economic Studies, 65, 453-473.

Francois, P. and Lloyd-Ellis, H. (2008). "*Implementation Cycles, Investment and Growth*," International Economic Review, 49(3), 901 - 942.

Franses, P. H. and Paap, R. (2004). "Quantitative Models in Marketing Research," Cambridge University Press: Cambridge.

Ghatak, M., Morelli, M. and Sjöström, T. (2007). "*Entrepreneurial talent, occupational choice, and trickle up policies*," Journal of Economic Theory, 137, 27 – 48.

Gilbert, C. L. (1986). "*Professor Hendry's Econometric Methodology*," Oxford Bulletin of Economics and Statistics, 48 (3), 283–307

Gleick, J. (1987). "Chaos: making a new science," Penguin Books: New York.

Golpe, A., (2009). "*Self-Employment and Business Cycles, University of Huelva, Spain*," PhD thesis.

Gourieroux, C., Monfort, A., and Trognon, A. (1984). "*Pseudo Maximum Likelihood Methods: Applications to Poisson Models*," Econometrica, 52(3), 701-720.

Gourieroux, C., Monfort, A., and Trognon, A. (1984). "*Pseudo Maximum Likelihood Methods: Theory*," Econometrica, 52(3), 681-700.

Greene, W. (2004). "*The behavior of the maximum likelihood estimator of limited dependent variable models in the presence of fixed effects*," Econometrics Journal, 7, 98–119.

Hausman, J. A. (1978). "*Specification Tests in Econometrics*," Econometrica, 46(6), 1251-1271.

Heij, C., Boer, P. de, Franses, P.H., Kloek, T. and Dijk, H.K. van (2004). "Econometric Methods with Applications in Business and Economics," Oxford University Press: New York.

Hodrick, R. J. and Prescott, E. C. (1980) "*Postwar U.S. Business Cycles: An Empirical Investigation*," Carnegie Mellon University discussion paper no. 451.

Hodrick, R. J. and Prescott, E. C. (1997). “*Postwar U.S. Business Cycles: An Empirical Investigation*,” Journal of Money, Credit and Banking, 29(1), 1-16

Juselius, K. (2009). "*Special Issue on Using Econometrics for Assessing Economic Models—An Introduction*," Economics: The Open-Access, Open-Assessment E-Journal, 3(28), 1-20.

Katz, E. (2001). "*Bias in Conditional and Unconditional Fixed Effects Logit Estimation*," Political Analysis, 9(4), 379-384.

Kim, C. J. and Nelson, C.R. (1999). “*Has the U.S. Economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle*,” The Review of Economics and Statistics, 81(4), 608-616.

Koellinger, P. D., Minniti, M. and Schade, C. (2008). “*‘‘I think I can, I think I can’’: Overconfidence and entrepreneurial behavior*," Journal of Economic Psychology, 28, 502–527.

Koellinger, P. D. and Thurik, A. R. (2012). “*Entrepreneurship and the Business Cycle*,” forthcoming in Review of Economics and Statistics. An earlier version would be Koellinger and Thurik (2009).

Koellinger, P. D. and Thurik, A. R. (2009). *“Entrepreneurship and the Business Cycle,”* Tinbergen Institute Discussion Paper TI09-032/3, Erasmus University Rotterdam

Krugman, P. (1979). "*A Model of Balance-of-Payments Crises*," Journal of Money, Credit and Banking, 11(3), 311-325.

Kydland, F. E. and Prescott, E. C. (1982). "*Time to Build and Aggregate Fluctuations*," Econometrica, 50(6), 1345-1370.

Lancaster, T. (2000). "*The incidental parameter problem since 1948*," Journal of Econometrics, 95, 391-413.

Lorenz, H. W. (1987). "*Strange Attractors in a Multisector Business Cycle*," Journal of Economic Behavior and Organization, 8, 397-411.

Maddala, G. S. (1983). “Limited Dependent and Qualitative Variables in Econometrics,” vol. 3 of Econometric Society Monographs, Cambridge University Press: Cambridge.

McKelvey, R. D. and Zavoina, W. (1975). “*A Statistical Model for the Analysis of Ordinal Level Dependent Variables*,” Journal of Mathematical Sociology, 4, 103–120.

Mundlak, Y. (1978). "*On the pooling of time series and cross sections*," Econometrica, 46(1), 69-85.

Pagan, A. R. (1987). “*Three Econometric Methodologies: A Critical Appraisal*,” Journal of Economic Surveys, 1, 3–24.

Potter, S. M. (1995). "*A Nonlinear Approach to US GNP*," Journal of Applied Econometrics, 10(2), 109-125.

Praag, M. van and Versloot, P.H., (2007). "*What is the value of entrepreneurship? A review of recent research*," Small Business Economics, 29, 351-382

Puu, T. and Sushko, I. (2004). *"A business cycle model with cubic nonlinearity*," Chaos, Solitons and Fractals, 19, 597–612.

Rampini, A. A. (2004). "*Entrepreneurial activity, risk and the business cycle*," Journal of Monetary Economics, 51, 555–573.

Ravn, M.O. and Uhlig, H. (1997). "*On Adjusting the HP-Filter for the Frequency of Observations*," CentER discussion paper no.50**.**

Ravn, M. O. and Uhlig, H. (2002). "*On adjusting the Hodrick-Prescott filter for the frequency of observations*," Review of Economics and Statistics, 84, 371-376.

Schumpeter, J.A. (1934). "The Theory of Economic Development", Cambridge, MA: Harvard University Press.

Self, S. G. and Liang, K. Y. (1987). "*Asymptotic Properties of Maximum Likelihood Estimators and Likelihood Ratio Tests Under Nonstandard Conditions*," Journal of the American Statistical Association, 82(398), 605-610.

Shleifer, A. (1986). "*Implementation Cycles*," The Journal of Political Economy, 94(6), 1163-1190.

Stel, A. van (2005). "*Compendia: Harmonizing Business Ownership Data Across Countries and Over Time*," International Entrepreneurship and Management Journal, 1(1), 105-123.

Stock, J. H. (1987). "*Measuring Business Cycle Time*," The Journal of Political Economy, 95(6), 1240-1261.

Stram, D. O. and Lee, J. W. (1994). "*Variance Components Testing in the Longitudinal Mixed Effects Model*," Biometrics, 50, 1171-1177.

Thurik, A.R. (2009), "*Entreprenomics: entrepreneurship, economic growth and policy*," Entrepreneurship, Growth and Public Policy, Z.J. Acs, D.B.Audretsch and R. Strom (eds), (Cambridge University Press, Cambridge), UK, 219-249.

Thurik, A. R., Carree, M. A., van Stel, A. and Audretsch, D.B. (2008). "*Does self-employment reduce unemployment?*" Journal of Business Venturing, 23, 673–686.

Vuong, Q., H. (1989). "*Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses*," Econometrica, 57(2), 307-333.

Wackerly, D. D., Mendenhall, W. III and Scheaffer, R. L. (2008). "Mathematical Statistics with Applications," Thomson: Toronto.

Wennekers, A.R.M., van Stel, A. J., Carree, M.A. and Thurik, A.R. (2010). “*The relation between entrepreneurship and economic development: is it U-shaped?*” Foundations and Trends in Entrepreneurship, 6(3), 167-237.

Wooldridgje, J. M. (1991). "*Specification testing and quasi-maximum likelihood estimation*," Journal of Econometrics, 48, 29-55.

Wooldridge, J. M. (2002). "Econometric Analysis of Cross Section and Panel Data," Cambridge: Massachusetts Institute of Technology.