Universitá degli Studi di Verona

Fostering Financial Stability: Three Essays on Banking, Potential Output and Employment

Dipartimento di Scienze Economiche

Scuola Superiore di Economia e Management
Dottorato di Ricerca in Economia e Finanza

Ciclo XXVIII
S.S.D. SECS-P/01

Coordinatore: Prof. Giam Pietro Cipriani
Tutor: Prof. Lorenzo Forni

Dottorando: Stefano Scalone

Stefano Scalone
Tesi di dottorato
Verona, Novembre 2017

Aknowledgements

First, I would like to thank my Supervisor Lorenzo Forni for the guidance and support he gave me in every stage of this PhD program.

I would then like to express my sincere gratitude to Professor Giam Pietro Cipriani, director of this PhD Program, and to the University of Verona for the quality of the didactics and the financial support.

My sincere thanks also goes to Professor Laurence J. Kotlikoff, host during my exchange at Boston University for his invaluable teachings and immense expertise.

I would then like to thank all my coauthors for their motivation and competence; it has been a great privilege to work with you all. A particular thank to Andrea, coauthor of Chapter two of this work for his support and infinite patience.

I thank all my PhD colleagues, and in particular my dear friends Monica and Claudio for having contributed to make the years of my PhD absolutely unforgettable, for their support and likewise their push for me to complete this program.

Finally, but my no means lasts, I would like to thank my partner Giorgia for the immense patience she showed and the invaluable help she gave me during all this time, and my parents, who encouraged and guided me throughout these years, supporting me in any possible way.

Dad, you have been and still are a great role-model, and your persistent push for me to finish this PhD has been crucial. It is for a big part your merit if I am here today.

Mum, you have been the most important person in my life and will always be, and I dedicate this work to you.

Contents

List of Figures

List of Tables

Summary

This work is devoted to the analysis of some of the fundamental pillars of economic stability. The first chapter deals with the banking sector, with the development of an innovative Early Warning System to timely identify cases of future financial distress, to allow the early intervention of the supervisor. The second chapter considers Gross Domestic Product and its potential, and discusses a new methodology for estimating potential output. The third and last chapter concludes by examinating the labor market, with the development of a new technique based on state-space modelling to estimate trend employment.

Chapter one presents an early warning model for predicting distress events tailored for small European banks. The novelty of the approach presented in this chapter is two-fold; first, the underlying approach broadens the usual definition of bank distress events, by basing it on the Bank Recovery and Resolution Directive (BRRD). In this way, the sample of distress events is significantly expanded compared to the most conventional approaches in the literature, despite considering a relatively short time span. This significantly improves the estimating power of the model. Second, the early warning system is modelled using the machin learning boosted decision-tree technique, á la Quinlan (via the C.50 Algorithm). We build a binary classification tree comprising bank-specific, banking-sector and macro-financial explanatory variables. The characteristics of the dataset we use are very peculiar: we use European Central Bank supervisory data on small European banks (the so called Less Significant Institutions, as defined by the ECB via its supervisory branch, the Single Supervisory Mechanism); our sample is strongly polarised towards one particular type of bank (simple retail) and a few jurisdictions (mostly Germany, Austria and Italy). For this reason, we first develop an initial model for the full sample of institutions, and then extend it in two direc-

tions. First, we subsample our data to re-run the decision tree on retail bank only. Second, starting from the baseline classification by Roengpitya et al. (2014) to classify institutions by business model, we build a decision tree to automatically sort banks into four lines of business: retail, wholesale, investment and custodians / asset managers.

The goal of chapter two is to build a "finance neutral" measure of potential output, developing and further extending the model by Borio et al. (2013). More proxies for the cycle are included in an augmented version of a state-space modelling framework, with a particular focus on liquidity (using short term debt as a proxy), and the analysis is extended to USA, UK, Spain, Italy, France, Austria, Netherlands and Switzerland. We confirm how a "finance neutral" measure of potential output is more reliable, even though the model results are quite tailored to the various jurisdictions: we find that different countries require different proxies, thus making it difficult to develop a one-fits-all model. Across-country robustness is in fact limited, and the best results are obtained by estimated an *ad hoc* model for each jurisdiction.

Chapter three proposes an innovative two-step approach to estimate the employment trend. The aim of this chapter is to incorporate demographic trends directly into the estimates, and get results that will incorporate the recent demographic dynamics that are strongly affecting the labor market: we here mean both purely demographic factors like population aging or higher life expectancy (in particular for males), and more labor market related trends like "baby boomers" reaching the age of retirement, decreasing the ratio between workers and pensioners or the steadily increasing level of women participation in the job market. We estimate the trend of employment using an augmented Kalman Filter procedure in a state-space framework, and conduct the estimation separately for each age cohort and gender (first step). We then aggregate the gender-cohort specific series (second step) to obtain the aggregate estimate of potential employment. By doing so, we are able to incorporate the specific trends of each cohort into the final estimate. The innovation is not only in the process, but also in the methodology: we augment the measurement equation of our state-space formulation to include some proxies for the financial cycle. We use this statistical refinement to obtain a cycle-free estimate, whose precision is signficantly higher with respect to traditional methodologies such as a HP or Kalman filter applied on the full sample of employment: this both reduces the indeterminacy of the estimation, and produces a more accurate quantification of the cyclical and trend components (e.g. by reducing the end-point problem). The improvements of this methodology are amplified in the latest years of our time series, those of the recent global economic crisis: a simple Kalman Filter applied on the full employment dataset tends to underestimate the fall in the cyclical component of employment with respect to our method. This result confirms how a methodology like ours is necessary to cope with the structural changes in labor force due to the recent demographic and labor market trends.

Chapter 1

A new approach to Early Warning Systems for smaller European banks

1.1 Introduction

Models for the early identification of bank financial distress represent a useful tool in the hands of the supervisor. They allow for timely intervention and in most cases the triggering of policy actions before the financial situation of an institution further deteriorates.

Existing early warning models are usually based on conventional modeling techniques such as multivariate logit models, and are calibrated using only a very small number of distress events.

With this work, we propose a two-fold innovation: first, creating a new definition for distress event based on the Bank Recovery and Resolution Directive (BRRD), we construct a sample of distress events over our dataset of small European banks that is significantly larger than any other work in the literature. Second, we propose a machine learning methodology to build a boosted decision three, which notably improves the predictive performance with respect to the most usual modeling techniques (we benchmark our decision tree with a logit estimation).

As said, our proposed approach in defining bank distress enlarges the usual sample size of distress events and therefore improves the learning of the model. We propose to classify banks as distressed based on the triggering of the BRRD's early interventions measures and on its criteria for categorizing banks as failing or likely to fail. Since this definition does not constitute the final stage of a bank's failure, the system will predict the *pre-failure* stage early enough to allow the supervisor to adopt preemptive measures to tackle the financial deterioration case.

This paper makes use of a decision tree model, a technique often applied in machine learning for classification problems, with the goal to construct a flexible and interpretable signalling tools for banking supervisors. The proposed early warning system (EWS) is able to predict individual bank distress events and identify which variables are significant to identify cases of financial deterioration.

This theoretical framework is applied to a unique dataset of more than 3,000 small European banks, the so called Less Significant Institutions (LSIs[1]). The EWS for LSIs (LSI-EWS) is built using three types of variables: bank specific ones, banking-sector indicators and country-level macro-financial variables.

The remainder of this paper is organised as follows: section 1.2 introduces the framework by analysing the European banking sector, section 1.3 summarises the past relevant literature on the topic, section 1.4 describes the dataset used for the estimation, section 1.5 presents the model and the main results, section 1.6 describes an extension of the model and section 1.7 concludes.

1.2 The European Banking System

The European banking system is extremely assorted and comprises institutions of different size, scope and business model, which range from big globally significant institutions (G-SIBs) to small local savings and cooperative banks. The banking sector is composed of around 120 banks[2] classified as *significant*, and supervised at the European level by the European Central Bank (through its supervisory branch, the Single Supervisory Mechanism, SSM), and more than 3,000 smaller institutions classified as *less significant* (LSIs from here on), the direct supervision

[1] As defined by the Single Supervisory Mechanism of the European Central Bank.

[2] Data source: EU Banking Supervision Website.

of which is left to the national competent authorities (but still conducted in close collaboration with the ECB).

The vast majority of these small institutions is of two types, *mutual savings* and *cooperative banks*. These types of banks originated in Europe between the late 18$^{\text{th}}$ and early 19$^{\text{th}}$ Centuries, with the goal of offering banking services to farmers, workers and small entrepreneurs, which at the time were facing extreme difficulties to access credit.

Even though these types of bank business models are extremely diffuse throughout Europe, the composition of the system of *less significant institutions* is unevenly distributed and notably concentrated in certain jurisdictions, namely Germany, Austria and Italy. The main reason behind this is historical, as even though most European countries have very deeply-rooted local systems of mutual savings and cooperative banks, in some jurisdictions these smaller institutions are unified in a single consortium or network (e.g. Credit Agricole in France or Rabobank in the Netherlands), and therefore are considered as *significant*.[3] As a consequence, the sample of European *less significant institutions* is strongly polarised towards a few countries, with Italy, Austria and Germany alone accounting for more than 80% of European LSIs.

This strong polarisation and the high absolute number of institutions raise some questions on how to timely identify cases of financial distress even before or as soon as they materialise, in order to give the supervisor enough time to intervene. This paper proposes an innovative method to identify cases of financial distress, by developing a boosted decision tree model based on European supervisory data, shaped in order to timely highlight - in the vast sample of Less Significant Institution - which banks are in distress and should therefore be monitored more closely.

[3]By the European supervisor, according to precise set of rules contained in the SSM "Guide to Banking Supervision:
https://www.bankingsupervision.europa.eu/ecb/pub/pdf/ssmguidebankingsupervision201411.en.pdf

1.3 Literature Review

1.3.1 Early Warning Systems

Economics is only one of the disciplines that makes use of early warning systems, as they are used in the most various subjects from disaster management for natural events, to medicine for timely identification of diseases, to the world of social media. Researchers in these subjects often borrow calculation and modelling techniques from other disciplines like physics and engineering, the same approach we use here with boosting decision trees and machine learning.

Our model is developed at the level of the single bank, an approach that differs with the general strand of literature: in economics, in fact, much more diffused are early warning systems to timely recognise signs of potential systemic risks, not only in the banking sector but at the country level. It is the case for example of Kaminsky (2000)'s seminal paper, developing an early warning model of financial vulnerabilities by creating some composite indicators of financial distress, based on the commonly shared hypothesis that economic distress and fragility of an economy are good predictors of financial crises.

The literature on single indicators of financial distress is vast, and often exceeds (or consciously diverges from) the scope of canonical early warning systems. A fundamental reference on the topic is Kaminsky and Reinhard (1999), who focus on currency crises. They adapt the methodology of Stock and Watson (1989) for leading indicators to develop a first step in the design of an early warning system that would help detecting a domestic financial crisis. The main signals are indeed currency and exchange rate expectations, but the authors underline how micro data on banking would be the natural complement of their analysis: they find that currency crises deepen banking crises, and analyse the deep interlinkages between these two sectors.

Focusing on models for banking crises, there is a tendency in the literature to identify indicators signalling banking problems, rather than building a fully specified model to predict bank distress. This is the case for example of Honohan (1997), with the author proposing a set of variables, each of which with a threshold that if exceeded can successfully predict a situation of bank distress. This apparently dogmatic approach however leaves room for the discretionary intervention of the supervisor, the *expert judgement* procedure that is extremely common in this strand of literature.

One of the first milestones in the literature on bank distress is represented by the pioneering work by Sinkey (1975), who adapted Altman (1968) to predict bank crises in a framework of multiple discriminant analysis. As outcome of the model, the variables indicated as good predictors for bank distress are asset composition, loan characteristics, capital adequacy, sources of revenues, efficiency and profitability.

Many of these variables correspond to the widely used CAMELS indicators, initially published by the Federal Financial Institutions Examination Council in 1979 and further developed by the Federal Reserve in 1995. This set of indicators represents a supervisory rating system to evaluate a bank's financial conditions and operations, and is widely used in the literature on banking distress identification. The indicators are:

- Capital adequacy,
- Asset quality,
- Management,
- Earnings,
- Liquidity,
- Sensitivity to market risk,

and represent a useful starting point for the variable selection of any banking model.

Not surprisingly, the interest in the literature on predicting bank distress events peaked after the 2007 financial crisis: Jin et al. (2011 and 2013) further developed the CAMELS approach by complementing the six indicators with data on banks' internal controls on risk-taking and audit quality variables, to find an improved predictive rate. Cole and White (2012) found that measures of commercial real estate investments are also relevant for predicting bank distress. We partly rely on this literature to construct the initial dataset for our boosted decision tree. We in fact build the model using three sets of variabels: bank specific indicators, banking-sector and country level macro-financial variables using the CAMELS approach as one of our starting points.

As the occurrence of a crisis - no matter how it is defined - can be easily described by a dummy variable, a common approach in the literature is to use logit/probit models. Thomson (1992) and Cole and Gunther (1998) estimate logit and probit models to show that vulnerability indicators covering the CAMELS dimensions are good predictors of banking failures. This is indeed a valid methodology, that we use to benchmark our estimations via the decision tree.

1.3.2 Decision trees

Our early warning system is based on a boosted decision tree methodology. The predictive performance of this technique is very high, both in and out of sample, as demonstrated by the data on accuracy presented in section 1.5; moreover, this methodology well handles missing values, a common issue in the early warning literature (Mitchell 1997), and one of the main flaws in our dataset. Finally, the boosted decision tree methodology is relatively transparent, and allows supervisors to interpret the output tree and understand which indicators affect bank distress, thus minimizing the risk of creating a *black box* model.

A decision tree is a classification technique commonly used in machine learning. The tree recursively identifies the significant indicators and their respective thresholds which best split the sample into the pre-determined classes (in our case distress and no distress).

More in details, a decision tree is a hierarchical model that identifies local regions (leaf nodes) via a sequence of recursive splits. Starting from the root of the tree, at each non-leaf node a test is run and a decision (one of the branches) is taken, depending on the outcome of the test. This process is repeated recursively, until a leaf node is reached; at that point the splitting stops, and the observation si labelled with the value of the leaf node.

The splitting procedure simply defines a region of the n-dimensional input space, where n is the number of distinct variables present in the tree; all the observations falling in the same localised region are given the same label by the model.

Among all possible splits, the best one is taken. The goodness of fit is measured in terms of *purity* and *entropy* (Quinlan, 1986). The purity of a node is a measure of the quality of the split: a node is pure if after the split, for all branches, all observations assigned to one branch belong to the same class. One possible way to measure the (im)purity of a node is *Entropy*:

$$H = -\sum_{i=1}^{n}(P_i log_2 P_i) \tag{1.1}$$

which is linked to the underlying probability of occurrence of value i: high entropy indicates that all classes are (nearly) equally likely (high impurity), while low entropy indicates that few classes are likely, and others are rarely observed (high purity).

The concept that guides the choice of the split is the maximization of the *information gain*, based on *conditional entropy*:

$$IG = H - (H_L * p_L + H_R * p_R) \tag{1.2}$$

where H is the *entropy* of the parent node, H_L is the entropy of the left node, and p_L is the probability that a random input is sent to the left node.

The final output of this recursive classification technique is a tree, illustrating a set of if-then rules (decision nodes) to reach a final decision on the classification (leaf nodes). In our case, for each bank, the classification starts from the root decision node, and based on predictors values create a path along the tree untila leaf node is reached, classifying if the bank is in distress or not.

1.4 Data, and the Sample of Distress Events

We use a dataset of supervisory data for around 3,000 European *less significant institutions.*

In order for an early warning system on bank distress to be useful in practice for the the supervisor and the policy maker, the recognition of the distress event must be timely enough to allow a buffer of time for intervention. If we consider the failure or liquidation of a bank as triggering event, as often defined in the literature, we lose the practical validity of the model, which would in turn be helpful only for ex post calibrations. Moreover, bank failures in Europe are relatively rare, this making the estimation of such an early warning system even more challenging. We therefore relax the traditional hypothesis of considering only bank crises or defaults as positive events in the sample (as e.g. done by Kaminsky and Reinhard 1999, who mark the beginning of a banking crisis by a bank run leading to closure, merging or take-over by the public sector of a bank, or large-scale government assistance), and instead consider all financial distress cases. By doing so, the sample of distress events significantly grows in size, allowing us to obtain precise estimations despite the short time horizon on which we span our model.

The large number of small institutions in Europe, together with the high quality of our supervisory data allows us to build a large dataset of distress events; we start by describing our relaxed hypothesis to identify bank distress events.

We use a mixed approach, and base our definition of distress on the Banking Recovery and Resolution Directive complemented by one of the four conventional types of financial distress in Betz et al. (2013), and end up with a database of more than 350 distress events throughout a sample of only six quarters. We consider a bank to be in financial distress if:

- It is deemed to be failing or likely to fail within the meaning of Article 32 of the BRRD. For categorizing a bank as failing or likely to fail, indicators assessing whether a bank has breached the minimum capital requirements or capital buffers are constructed;

- It meets the conditions for early intervention pursuant to Article 27 of the BRRD. The triggers used to meet the conditions of early interventions consist of indicators for assessing if a bank is close to breaching minimum capital requirements;

- It is placed under special administration and/or is appointed of a temporary administrator pursuant to Article 29 of the BRRD;

- There is a rapid and significant deterioration of its financial situation according to Article 96 of the Framework Regulation. This is based on expert judgement by national central banks and in-house qualitative and data;
- One of the four types of conventional bank distress events proposed by Betz et al (2013) (i.e. bankruptcies, liquidations, state interventions and forced mergers) is met.

1.5 Methodology

1.5.1 Data pre-processing

Data pre-processing steps are required to ensure that unreliable and noisy data as well as irrelevant and redundant information is eliminated prior to the modelling phase. As such, the final training dataset used for the analysis is of high quality, thus increasing the efficiency and performance of the final model. This represents a key step in our process, as we start from the extremely vast dataset of supervisory reporting consisting of more than 3,000 variables, and end up with a final sample of only 12.

We start by cleaning the data, in order to eliminate incomplete or uninformative data; many banks do not report some data points, mostly because they are not relevant for them or do not apply to their business model. We therefore eliminate variables for which the majority of values is missing, or where the variance across the sample is close to zero. This first step already reduces the number of total indicators to less than 500, this giving a hint of the importance of the data pre-processing procedure.

The second step is a simple transformation of the data, with the goal of increasing consistency and comparability across institutions. More than one accounting standard coexists in Europe, this complicating the job of the supervisor; we apply a transformation technique, based on a mapping of these heterogeneous data points to make different data sources somehow coherent. In a following stage, we normalise our variables through the creation of ratios in order to increase comparability.[4]

We remove explanatory variables which are too highly correlated, using a simple threshold of 0.9 and select the final set of indicators based on their ability to predict distress. Variables are ranked according to their importance, captured by the individual Area under the Receiver Operating Characteristic curve (AUC) for each indicator; using this technique, we select the top 100 variables in terms of predictive performance which we use as starting point for the boosted decision tree.

1.5.2 Model and results

We employ Quinlan's C5.0 algoritm to build the classification tree model. The C5.0 algoritm is one of the most commonly used, as it is relatively fast and accurate, as well as efficient in handling missing data and removing unhelpful attributes.[5]

In training the model, we select a number of specification options:

- We impose a relatively short prediction horizon (1-3 months ahead of distress as starting point), given the short term (<1 year) scope of this EWS. By considering pre-default events as target variable, we ensure that the system has a forward looking perspective;
- We impose asymmetric misclassification costs when assessing the performance of explanatory variables: we consider Type I errors (missing a distress events) to be twice as costly as Type II errors (issuing a false alarm). In principle, this assumes that when faced with a tradeoff of issuing more false alarms or missing a distress event, the policymaker would take a conservative stance and choose the former;
- To increase the robustness of simple decision trees (which is relatively low, as udnerlined by Alessi and Detken (2014), who use a Random Forest method

[4]The full list of variables and the mapping of different accounting standards is available in Annex 1.9.

[5]For a literature review of Data Mining Algoritms see Wu et al. (2008). The relative *R* environment used in this paper refers to Kuhn et al. (2015).

to overcome the problem), we employ a boosting technique á la Freund et al. (1999) to identify which variables to include in the final version of the tree.[6]

- As there is no univocal rule to choose one particular tree among the estimated ones, we use the boosting technique to simulate the creation of a large number of trees, and select the 20 variables that rank highest as of importance (measured in terms of how often they appear in the trials).

- We complement the variable selection with both expert judgement and quantitative measures: in particular, for evaluating the performance of the model, we rely on the area under the Receiver-Operating-Characteristic curve (AUC) and Cohen's kappa statistic, both standard measures of accuracy in the early warning system literature (e.g. Peltonen at al. (2015)).

The final tree is composed of 19 nodes, covering 12 different explanatory variables, and is represented in Figure 1.2.[7]

The variables included in the model are:

- Adjusted profitability;
- Non-performing loans (NPL) ratio;
- Non-performing loans coverage ratio;
- Deficit-to-GDP ratio;
- GDP growth;
- Liquidity coverage ratio (LCR);
- Leverage ratio;
- Equity exposures;

[6]Boosting is a technique for generating and combining multiple classifiers to improve the predictive accuracy of the model. Instead of using a single tree, n separate decision trees (trials) are grown and combined to make predictions. The error rate of the boosted classifier is often substantially lower than that of single trees.

[7]Please note that the trees represented in this version of the paper are somehow anonymised, i.e. without the precise splitting thresholds of each node.

- Exposures in default;
- Two proxies for market risk;[8]
- Membership in an institutional protection scheme (IPS),

The indicator of the parent node is profitability, adjusted for the different accounting standards of the banks in the sample. The node splits the banks between profit (right branch) and loss (left branch) making. The remaining variables are a mix of macro-economic indicators (deficit-to-GDP ratio, and real GDP growth) and banking indicators covering the most important risks: credit (non-performing loans ratio, non-performing loans coverage ratio, exposure in default), liquidity (liquidity coverage ratio), market (captured by the sum of trading financial assets and financial liabilities held for trading over total asset and net gains on financial assets and liabilities held for trading over total operating income), capital (leverage ratio and equity exposure), together with the qualitative information of whether a bank is member of an institutional protection scheme (IPS).

In the framework of supervised learning, the role of our tree is not only to find an efficient method to identify which banks are in financial distress, but also to suggest which variables are significant and how they model the distress. In this perspective, it is interesting to analyse the main paths through the tree: if a bank is making profits (parent node to the right), profitability is likely to not be an issue, and the model suggests to investigate credit risk (first node to the right is the NPL ratio). If credit risk is deemed as material, i.e. if the bank has a high level of non-performing loans, the model moves to analyse whether these NPLs are covered by sufficient allowances. On the other hand, for banks with low non-performing loans (NPL ratio node to the left), market risk becomes relevant in capturing distress: this is intuitive, as banks strongly relying on income from market activities are subject to higher volatility and potential distress, especially in countries with fragile economic fundamentals (high level of deficit over GDP ratio).

On the opposite side of the tree, if a bank is deeply unprofitable (parent node to the left, first node to the left), if it is operating in a country with a low growth

[8]The sum of trading financial assets and financial liabilities held for trading over total assets, and net gains on financial assets and liabilities held for trading over total operating income.

of GDP it is automatically labelled as being in distress. When instead the bank has relatively high equity exposures combined with a weak LCR ratio, then the distress will be determined by the eventual membership in an IPS; Institution Protecting Schemes in fact protect banks from financial distress, therefore making member institutions less vulnerable. Finally, for moderately unprofitable banks, the economic conditions of the country in which they operate (proxied by deficit over GDP) and the leverage ratio indicate whether a bank is in distress or not.

The predictive performance of the model is very high, both in and out of sample. As depicted in Table 1.1, in the training data the true positive rate is 0.89, while the false positive rate (Type II error) and false negative rate (Type I error) equal 0.01 and 0.11, respectively. The AUC is 0.95, much closer to the unity value of a perfect classifier, than to the 0.5 value of a purely random one. The Cohens kappa statistic is also high, at a value of 0.89. The out of sample performance (25% of the initial observations) is also satisfactory, with an AUC of 0.92 and a Cohens kappa statistics of 0.80. Type I and II errors are therefore comparable to the in-sample error rates, and remain at adequately low levels.

Table 1.1: Validation Results Full Sample

Measures	In Sample (train)	Out of Sample (test)
Type I error rate	0.01	0.03
Type II error rate	0.11	0.10
AUC	0.95	0.92
Cohen's Kappa	0.89	0.80

It is useful to remark how the model is estimated by introducing an unbalance between level-1 and level-2 errors: false positives are in fact considered to be less problematic than missed distress events, and therefore weight half.

We benchmark the results of our boosted decision tree model with a simple Logit regression, an approach typical in the literature of predictions on a categorical target variable (in our case the binary event of a bank financial distress).

We automatically select the variable via a LASSO regression (Tibshirani, 1996) in order to prevent overfitting, and for the sample preparation follow the same steps of the decision tree.

The results are still quite accurate, but the Logit model misses many more distress events than the decision tree. The Logit is in fact significantly more sensitive to missing values, and therefore fails to detect distress events relative to institutions reporting a sufficient number of NAs, unlike the decision tree.

Table 1.2: Validation Results Logit Model

Measures	In Sample (train)	Out of Sample (test)
Type I error rate	0.01	0.02
Type II error rate	0.18	0.19
AUC	0.95	0.90
Accuracy	0.97	0.97

1.6 Model Extension

1.6.1 Composition of the sample

As explained in details in section 1.2, the European banking system, (and in particular the sample of the so called *less significant institutions*, on which we base our sample and create the model) is very peculiar. The vast majority of them are simple retail banks, strongly rooted in the territory, mostly taking deposits from local customers and giving mortgages to families and loans to SMEs. Regarding the jurisdiction in which they operate, more than half of the banks in our sample are German. Germany and Austria alone account for two thirds of the whole dataset. The sample is therefore extremely polarised towards one business model and two jurisdictions.

On the other hand, the rest of the dataset is composed of banks that are extremely different from these simple and relatively small retail lenders, and very heterogeneous among them: it comprehends institutions from 17 other jurisdictions, and with business models ranging from investment banks, custodians, payment system banks, specialised finance institutions (e.g. banks focused almost entirely on car financing operations).

This has two straightforward consequences: first, the strong heterogeneity in the sample implies that it is mostly retail lenders that shape the model, which therefore performs worse on the other business models (it misses 9% of distress events for retail lenders, and 25% for other business models). Second, the remaining part of the sample is not only small, but also extremely heterogeneous in its composition: we have more than 10 different business models in the less than 10% of banks that are not simple retail.

As a consequence, we here develop a model for a sample including only retail lenders. We therefore exclude all other business models, and try to derive some new insights.

1.6.2 A tree for retail lenders

We here describe how the model changes, if we only consider retail lenders in the sample.

These banks represent mode than 90% of the initial sample, and most of them are quite homogeneous in terms of activities: they are focused on lending to the retail sector, with their lending portfolio usually strongly polarised towards residential real estate lending, sometimes complemented by lending to the so called SMEs (small or medium size enterprises). The main source of funding is often represented by the deposits of that same base of retail clients, and a variable part from the wholesale sector. Their main source of income is usually interest income, coming from the spread between the rate on loans and cost of funding, together with income from fees and commissions, mostly stemming from fees coming from their lending activities.

The result of this *ad hoc* analysis can be found in Figure 1.3, in the form of a boosted decision tree model.

It appears quite clearly that the tree for retail lenders is significantly simpler than the one run over the whole sample. This is because part of the sample variation was manually excluded, and throughout the new sample banks tend to be similar and share many features.

The variables included in the model are a subsample of those in the tree from the previous section, and include:

- Adjusted profitability;
- Bank leverage;
- NPL Ratio;
- Allowances for NPLs;
- Exposure in default;
- A proxy for trading income[9].

Once again, adjusted profitability represents the parent node. To the left, profitability is still the best predictor and splits institutions into *poorly* - which are immediatly labelled as distressed - and *extremely poorly* profitable - for which more splits are needed. We interpret this apparently counterintuitive result as follows: first, profitability is taken in its absolute value and not normalised terms (as e.g. return on equity or return on assets): the size of the bank could therefore play a role in this split[10]. Second, our hypothesis is that the model is able to capture the fact that a strongly negative profitability could be *artificial*, as in the example of a bank with a strong negative one-off effect (take for example the cases of Unicredit or Deutsche Bank in 2016), but relatively healthy otherwise.

Following this path, the model in fact moves to analyse trading income and leverage ratio, as if it focused on understanding whether the strongly negative profitability came from a bank with high variability in the profits. This is accounted together with equity exposure and NPL ratio that, if high, represents a strong sign of financial distress for a bank with low profitability.

On the other side, the tree broadly follows the model developed in section 1.5.2: once profitability is high, the most relevant variable is the NPL ratio. If this is high, then we look at the coverage on these loans, and so on (continuing with equity exposure and market risk).

Institutions are however split unevenly across the model: one leaf nodes contain a vast majority of banks. For non-distressed institutions, the node representing

[9] In particular, the ratio between net gains from financial assets and liabilities held for trading and total operating income.

[10] We extensively tried models with normalised indicators of profitability, but found that the absolute adjusted level was steadily the best predictor.

medium to high adjusted profitability and low level of non-performing loans contains a large part of the sample. This is a result that is in line with our expectations, as these are two clear signs of a bank with a healthy business.

Table 1.3: Validation Results Retail Lenders

Measures	In Sample (train)	Out of Sample (test)
Type I error rate	0.01	0.01
Type II error rate	0.10	0.15
AUC	0.95	0.94
Cohen's Kappa	0.91	0.86

1.7 Conclusions

This paper develops an innovative model to identify cases of bank financial distress, using a subsample of 3,000 small European institutions[11] for a time period of six quarters (between 2014 and 2016).

We build a sample of distress cases based on the BRRD regulation, to early detect future cases of financial deterioration rather than simply referring to banks that are already in or close to default. With a broad definition of financial deterioration, our sample of distress events is significantly larger than any other work in the literature, despite a relatively short time series of data.

On this sample we construct a boosted decision tree model, which accurately classifies banks into distressed or non-distressed; the prediction horizon of the model is one-three months, a time span that would in our view give the supervisor enough time to trigger supervisory action.

We find that the predictive power of our model is extremely high, and the decision tree steadily outperforms the Logit approach, the most widely used methodology to predict binary classifications which we use as benchmark.

As a final remark, further extensions of this work should go in the direction of increasing the prediction horizon (currently 1-3 months), and could include the

[11]Excluding the so called Significant Institutions, as defined by the SSM.

development of an *ad hoc* tree for each business model. Unfortunately, the data does not allow us to do this, yet, as given the polarisation of the dataset towards retail lenders, the sample of remaining business models is simply not numerous enough to allow for a proper estimation.

Moreover, once the time series of the model will be long enough, it will be interesting to analyse the changes in the business models of the banks; this could be achieved in two ways: first, by re-estimating the model to understand how the environment changed, and what new variables contribute to describing the business of an institution; second, by conducting a case-by-case analysis for institutions for which the model changes classification: in this way, it would be possible to get some insights on how the changes in monetary policy and the economic environment influence the behaviour of banks (think for example of the current prolonged period of low interest rates, which might be pushing retail banks that see their interest margins eroded into finding new sources of income).

Finally, we plan to develop a new model able to identify the severity of bank distress events. In order to do so, we will buikd a multi-class target model to classify distresses as Mild, Moderate and Severe. Such classification would be important for supervisors in order to efficiently allocate resources and improve financial surveillance.

1.8 Appendix A: A Tree for Business Model Classification

As annex to this chapter we present a first further extension of this work, the development of a boosted decision tree to classify banks according to their business model.

In order to isolate retail lenders from the rest of the sample, we divide the institutions in the dataset into different business models. Following Roengpitya et al. (2014), we split banks into four business models: retail lenders, wholesale lenders and capital market-oriented institutions (then further split into investment banks and custodians/asset managers). The business lines of retail lenders have already been discussed in details; wholesale lenders mainly differ from the retail ones for their source of funding, mainly corporate; capital market-oriented banks represent in our case a residual and quite heterogeneous category: this group includes not only pure investment banks, but also other institutions whose main income sources are fees or trading.

Rather than simply splitting the model into the various bank types, once again we let the data speak and develop a boosted decision tree model, to understand which variables are relevant and which are the business lines that help sort the institutions.

The variable selection process is similar to the one employed in the decision tree developed in section 1.5.2, but the initial sample of variables is much smaller. In order to understand the business model of an institution, from the initial sample of variables we drop the indicators that Roengpitya et al. (2014) call *outcome variables*, i.e. those whose outcome is independent of the institutions (e.g. all variables from the statement of profit and loss), and focus on those that derive from precise strategic choices of the bank (like the composition of the balance sheet). For the analysis of the full set of indicators considered in the dataset refer to Appendix 1.9; the final set of variables included in the model comprises: complexity (proxied by the percentage of credit risk exposures over total risk exposures), total deposits (over total assets), total household deposits (over TA), total loans (over TA), impairments (over TA), total costs (over TA), and financial liabilities held for trading.

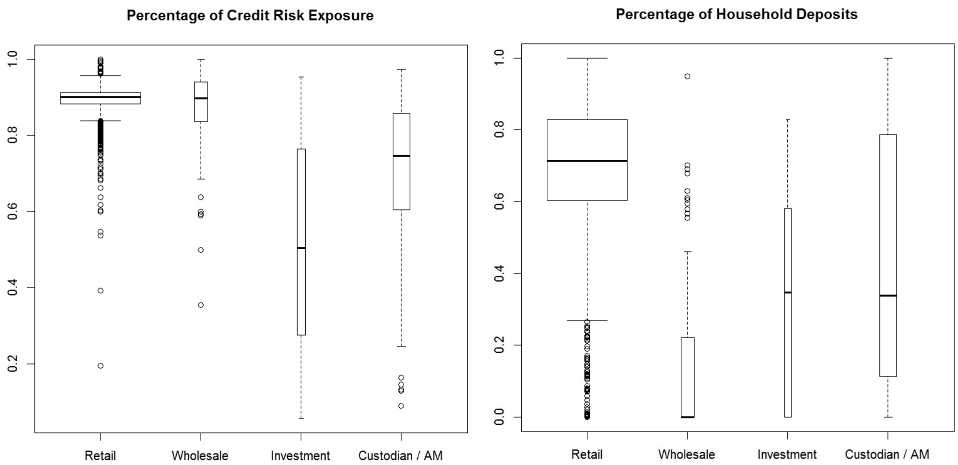
Table 1.4: Validation Results Retail Lenders

Measures	In Sample (train)	Out of Sample (test)
Accuracy	0.98	0.94
Multiclass AUC	0.84	0.77
Cohen's Kappa	0.91	0.64

The performances of this model are somehow lower with respect to the ones of the previous sections, as not only the initial set of variables is limited (as said, we only include variables that are clear strategic choices of the bank, and exclude all the others), but also the classification is performed into four classes instead of two. The results are still satisfying, and we briefly describe them below.

By analysing the composition of the decision tree, we can trace the main features characterising the business model of a bank: credit risk as a minor share of the business is a strong indicator of a capital market-oriented institution, while retail banks tend to be on the right side of the tree, among the *less complicated* ones.

Figure 1.1: Distribution of Credit Risk Exposures and Household Deposits



The vast majority of Retail Lenders are captured by the node to the far right, i.e. relatively low complexity and high household deposits. This node identifies very precisely the business model of retail lenders, four fifths of which are classified here. As a consequence, the rest of the nodes containing Retail Lenders (in light green in Figure 1.4) include a rather small number of banks with peculiar businesses (the only other node with a relatively high number of Retail Lenders is that with medium household deposits, a high level of total costs and low trading income).

Wholesale lenders tend to be relatively more complex (so with a significant share of risk not coming from credit) and small in size, with a medium to low share of household deposits (from which a low level of non-performing loans), and a more efficient cost structure.

With such a small number of variabels analysed, the difference between Investment and Custodians / AMs can be subtle, and at times even the model fails to detect the distinction between the two business models. The majority of misclassifications comes in fact from here. However, Investment banks in our sample are complex, do not raise much funding through household deposits (and therefore generally do not face a high level of NPLs) and are relatively small in size (our sample is limited to European banks below 30 bln € of total assets).

Similarly, Custodians and Asset managers have a low level of household deposits (and of NPLs), have a high level of own funds and are complex (generally more than Investment banks; the nodes indicating Investment institutions in the right part of the tree are quite small in number) as they tend to get a significant part of their income from trading.

1.9 Appendix B: List of Variables Used in the Model

1.9.1 Full Sample Decision Tree

- Adjusted profitability: {F02.00; R670; C010} + (flow from hidden reserves, internal data reference) + (funds from general banking risks, internal data reference);
- Bank leverage: {C45.00.b; R180; C040};
- Non-performing loans (NPL) ratio: ({F18.00.a; R070 + R250; C060} - {F18.00.a; R100 + R280; C060}) / ({F18.00.a; R070 + R250; C010} - {F18.00.a; R100 + R280; C010});
- Deficit-to-GDP ratio: ECB Statistical Data Warehouse (SDW);
- Real GDP growth: ECB Statistical Data Warehouse (SDW);
- Liquidity coverage ratio (LCR): internal data reference;
- Equity exposures: (C07.00a; R070; C010; 016 + C10.01.a; R020 + R050; C060) / (C07.00a; R070; C010; 001 + C08.01a; R020; R020; 001 + C08.01a; R020; C020; 002);

- Allowances for NPLs: ({F18.00.B; R070+R250; C150} - {F18.00.B; R100 + R280; C150}) / ({F18.00.A; R070 + R250; C060} + {F18.00.A; R100 + R280; C060});
- Exposure in default: {C07.00.a; R010; C010} + {C08.02; C020};
- Net gains from financial assets and liabilities held for trading over total operating income: (F02.00; R280; C010 + F02.00; R285; C010) / (F02.00; R355; C010);
- Sum of financial assets and liabilities HFT over total assets: (F01.01; R050; C010 + F01.01; R091; C010 + F01.02; R010; C010 + F01.02; R061; C010) / (F01.01; R380; C010);
- Membership in an institutional protection scheme (IPS): qualitative variable.

The list of variables is not provided for the retail lenders tree, as it is a subsample of the above.

1.9.2 Business Model Classification Decision Tree

When trying to identify the business model of an institution, we look at variables from all the following dimensions:

- Size: total assets as {F01.01; R380; C010};
- Complexity: credit RWA over total RWA as {C02.00; R040; C010} / {C02.00; R010; C010};
- Trading: financial liabilities held for trading over total liabilities as: {F0102; R010; C010} / {F0102; R300; C010}
- Cost structure: total costs over total assets as: {F0200; R360 + R390; C010} / {F01.01; R380; C010}
- Own funds: earnings over own funds as {C0100; R130; C010} / {C0100; R010; C010};

- Liabilities:
  - Total deposits over total assets as: {F0801a; R050; C010 + C020 + C030} / {F01.01; R380; C010}
  - Household deposits over total deposits as: {F0801a; R310; C010 + C020 + C030} / {F0801a; R050; C010 + C020 + C030}
- Exposures: total loans non performing over total loans as: {F1800a; R070 + R250; C060} / {F1800; R070; C010}.

Bibliography

[1] Alessi L. and Detken C. , 2014, *Identifying Excessive Credit Growth and Leverage*, ECB Working Paper n.1723.

[2] Alpaydin E. I., 2010, *Introduction to Machine Learning*, 2^{nd} Edition, The MIT Press.

[3] Altman E. I., 1968, *Financial Ratios, Discrimimant Analysis and the Prediction of Corporate Bankruptcy*, The Journal of Finance, 23, 589-609.

[4] Altman E. I., Marco G., and Varetto F., 1994 , *Corporate Distress Diagnosis: Comparisons Using Linear Discriminate Analysis and Neural Networks*, Journal of Banking and Finance, 18, 505-529.

[5] Atiya A., 2001, *Bankruptcy Prediction for Credit Risk using Neural Networks: A Survey and new Results*, Transactions on Neural Networks, 12(4), 929-935.

[6] Betz F., Opric S., Peltonen T. A., and Sarlin P., 2014, *Predicting Distress in European Banks*, Journal of Banking and Finance, 45, 225-241.

[7] Boritz, J. E. and D. B. Kennedy, 1995, *Effectiveness of Neural Network Types for Prediction of Business Failure*, Expert Systems with Applications, 9(4), 95-112.

[8] Cole, R. A., and J. Gunther, 1998, *Predicting Bank Failures: A Comparison of on- and off- site Monitoring Systems*, Journal of Financial Services Research 13(2), 103-117.

[9] Cole, R. A. and L. J. White, 2012, *Deja Vu all over again: The Causes of U.S. Commercial Bank Failures this Time Around*, Journal of Financial Services Research 42(1-2), 5-29.

[10] , Flannery M. J., 1998, *Using Market Information in Prudential Bank Supervision: A Review of the U.S. Empirical Evidence*, Journal of Money, Credit and Banking, 30(3), 273-305.

[11] Freund, Y., Schapire, R., and Abe, N., 1999, *A Short Introduction to Boosting*, Journal-Japanese Society For Artificial Intelligence, 14, 771-780, n.1612.

[12] Frydman H., Altman E. I. and Kao D. L., 1985, *Introduction Recursive Partitioning for Financial Classification: The Case of Financial Distress*, Journal of Finance, 40(1), 269-291.

[13] Gaytn A. and Johnson C. A., 2002, *A Review of the Literature on Early Warning Systems for Banking Crises*, Central Bank of Chile Working Papers n.183.

[14] Gepp A., Kumar K. and Bhattacharya S., 2010, *Business Failure Prediction using Decision Trees*, Journal of Forecasting, 29(6), 536-555.

[15] Gepp A. and Kumar K., 2015, *Predicting Financial Distress: a Comparison of Survival Analysis and Decision Tree Techniques*, 11^{th} International Conference on Data Mining and Warehousing.

[16] Gonzlez-Hermosillo M. B., 1999, *Determinants of ex-ante Banking System Distress: a Macro-micro Empirical Exploration of some Recent Episodes*, International Monetary Fund.

[17] Honohan P., 1997, *Banking System Failures in Developing and Transition Countries: Diagnosis and Predictions*, BIS Working Papers n.39.

[18] Jin, J., K. Kanagaretnam, and G. Lobo, 2011, *Ability of Accounting and Audit Quality Variables to Predict Bank Failure During the Financial Crisis*, Journal of Banking & Finance 35(11), 2811-2819.

[19] Jin, J., K. Kanagaretnam, G. Lobo, and R. Mathieu, 2013, *Impact of FDICIA Internal Controls on Bank Risk Taking*, Journal of Banking & Finance 37(2), 614-624.

[20] Joos P., Vanhoof K., Ooghe H. and Sierens N., 1998, *Credit Classification: a Comparison of LOGIT Models and Decision Trees*, Proceeding Notes of the Workshop on APplication of Machine Learning and Data Mining in Finance, 59-72.

[21] Kaminsky G. L., 1999, *Currency and Banking Crises: the Early Warning of Distress*, IMF Working Paper, WP/99/178.

[22] Kaminsky G. L. and Reinhart C. M., 1999, *The Twin Crises: the Causes of Banking and Balance-of-Payments Problems*, The American Economic Review, v.89 n.3, 473-500.

bibitemkohavi:02 Kohavi R. and Quinlan R., *Decision Tree Discovery*, in Handbook of Data Mining and Knowledge Discovery, 267-276, Klosgen & Zytkow, Oxford University Press, Oxford.

[23] Kuhn M., Weston S., Coulter N., and Culp M., 2015, *C5.0 Decision Trees and Rule-Based Models*.

[24] Meyer P. A. and Pifer H. W., 1970, *Prediction of Bank Failures*, Journal of Finance, 25(4), 853-868.

[25] Mitchell T., 1997, *Machine Learning*, 1st edition, McGraw-Hill, New York.

[26] Ng G. S., Quek C. and Jiang H., 2008, *FCMAC-EWS: A Bank Failure Early Warning System based on a Novel Localized pattern Learning and Semantically Associative Fuzzy Neural Network*, Expert Systems with Applications, 34(2), 989-1003.

[27] Peltonen, T. A., Piloiu, A., and Sarlin, P., 2015, *Network Linkages to Predict Bank Distress*, ECB Working Paper Series, n.1828.

[28] Quinlan J.R., 1986, *Induction of Decision Trees*, Machine learning 1, 81-106.

[29] Roengpitya R., Tarashev N. and Tsatsaronis K., 2014, *Bank Business Models*, BIS Quarterly Review, December 2014.

[30] Sinkey, J. F., 1975, *A Multivariate Statistical Analysis of the Characteristics of Problem Banks*,The Journal of Finance, 30(1), 21-36.

[31] SSM, 2014, *Guide to Banking Supervision*.

[32] Stock J. H. and Watson M., 1989, *New Indices of Coincident and Leading Economic Indicators*, National Bureau of Economic Research Macroeconomics Annual 1989, NBER.

[33] Tibshirani R., 1996, *Regression Shrinkage and Selection via the LASSO*, Journal of the Royal Statistical Society Series B, 58(1), 267-288.

[34] Thomson, J. B., 1991, *Predicting Bank Failures in the 1980s*, Economic Review-Federal Reserve Bank of Cleveland, 27(1), 9.

[35] Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., *et alii*, 2008, *Top 10 Algorithms in Data Mining*, Knowledge and Information Systems, 14(1), 1-37.

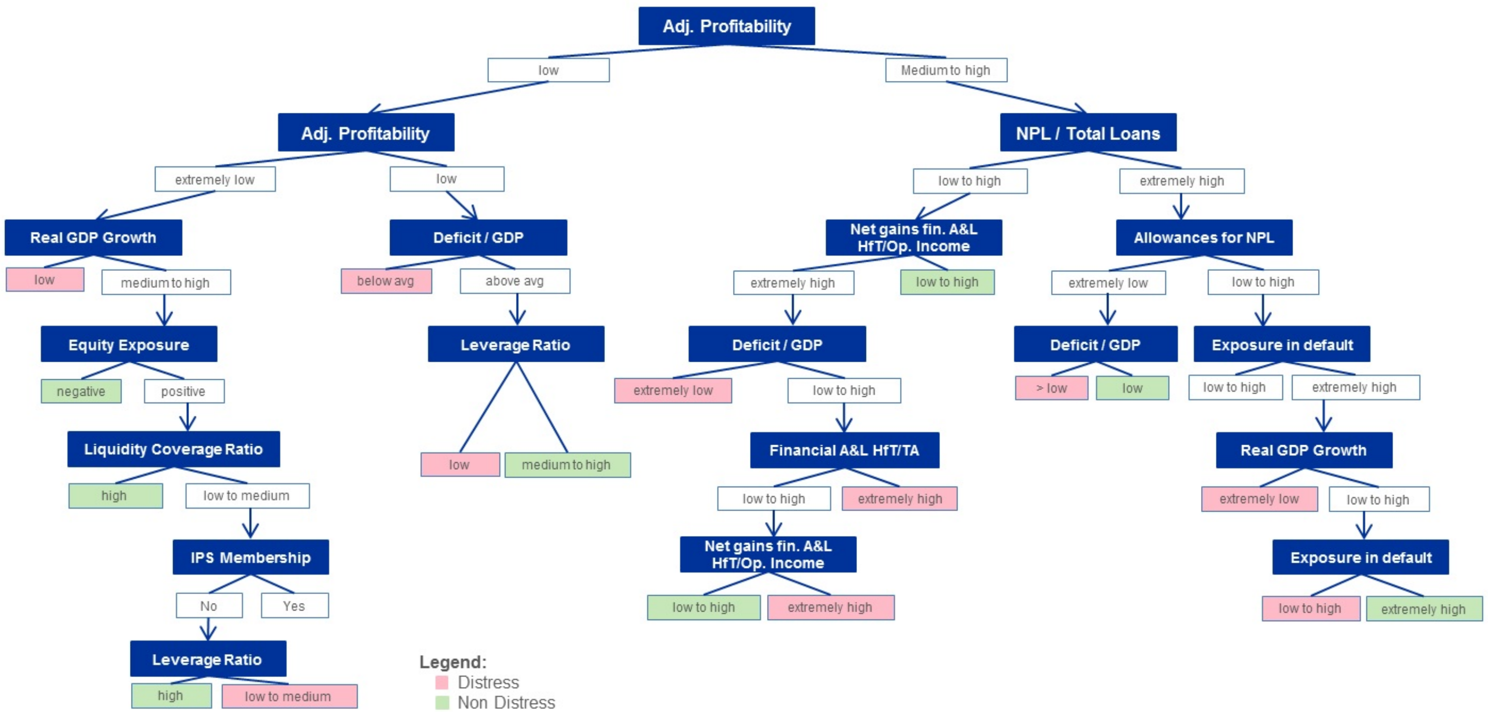


Figure 1.2: Boosted decision tree for full sample

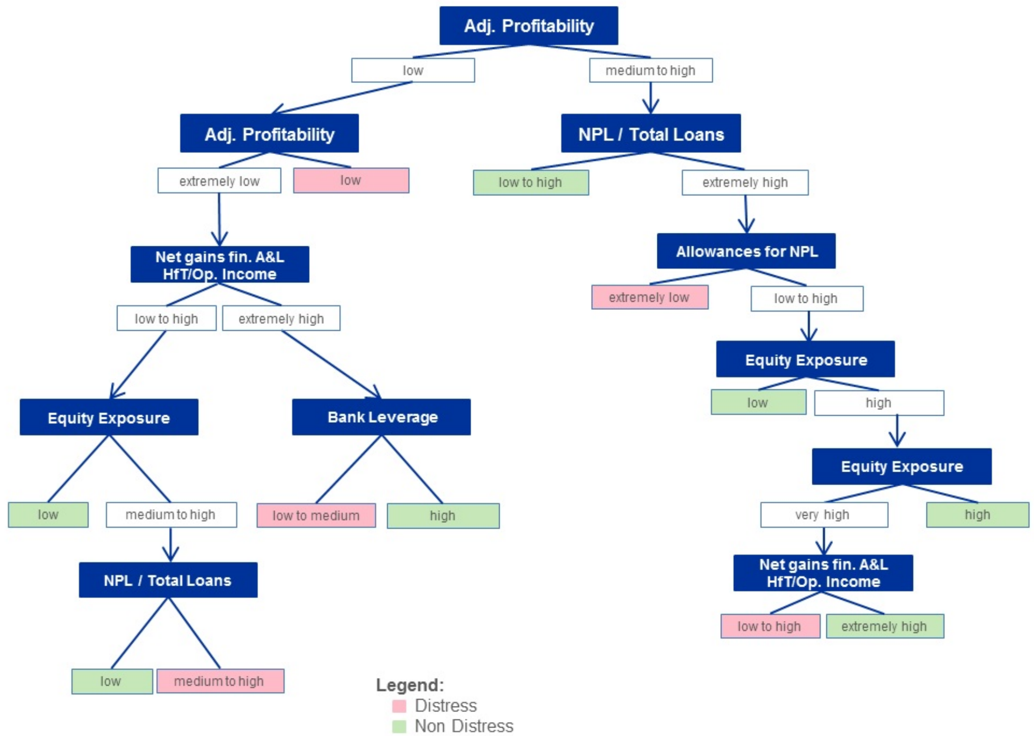


Figure 1.3: Boosted decision tree for subsample of Retail Lenders

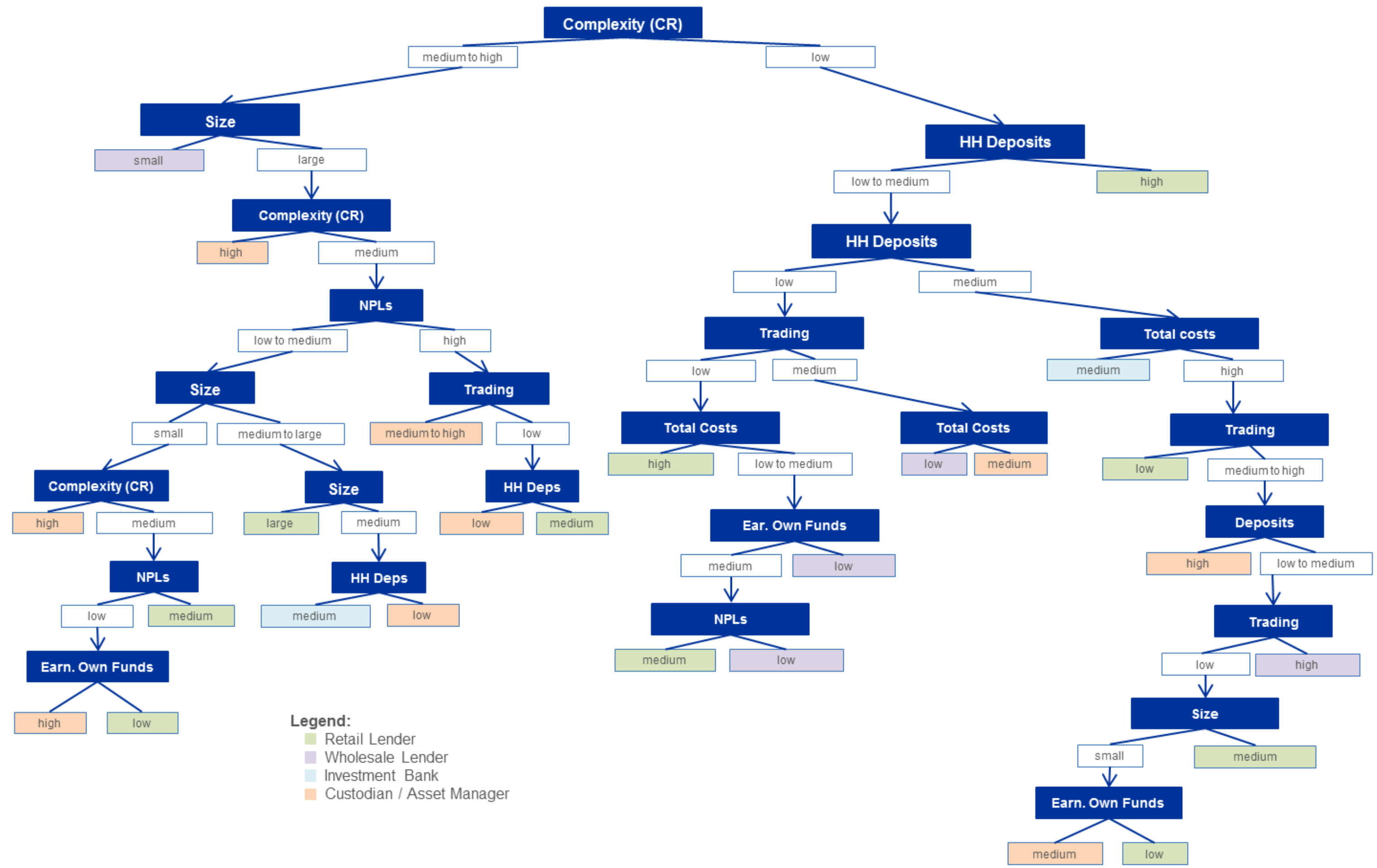


Figure 1.4: Boosted decision tree for Business Model Classification

Chapter 2

Embedding Liquidity Information in Estimating Potential Output

2.1 Introduction

The output gap is a fundamental concept in modern macroeconomics. Its relation with inflation is embedded in one of the three equations of the so-called New Consensus Macroeconomics, while governments and Central Banks make many policy decisions based on the deviations of actual output from its potential. While in the long run a faster growth of potential output corresponds to a growing path of actual output, in the short run policy makers have to analyze whether the output gap is due to desired responses of the economy to shocks, or to undesirable deviations from the optimal path. The issue of how to measure and estimate the output gap is therefore of primary importance.

Despite having a relatively simple definition (the gap is the difference between actual and potential output), there are many difficulties that impede its estimation: first and foremost, potential output is not observable, not even *ex post*. The debate on how to calculate it regards both the statistical and the conceptual level: from the one side researchers debate on which is the best way to estimate potential, from the other some argue that more information should be included in the estimation process.

In this work, we first describe the main issues in the estimation process, and then - taking Borio et al. (2013) as a reference - study how to further improve the output gap estimation methods. In a state-space modelling framework, in which the state equation is augmented to include some proxies for the financial cycle, we build a "finance neutral" version of potential output which is more robust and reliable than some of the most typical estimation methodologies in the literature. We however find that our model differs across jurisdictions, and obtain the best results when tailoring a "country specific" version of potential output.

2.2 The Crisis of the Phillips Curve and the Financial Cycle

The output gap appears to be a simple concept: it is the difference between actual and potential output. It assumes positive values when periods of strong demand drag the rate of utilization of capacity up, and falls below zero when demand decreases below the level which could be potentially achieved with the existing production capacity.

Calculating potential growth is a primary issue for policy makers: they make policy decisions based on these estimates. Estimates of the output gap are used for macroeconomic forecasts, provide a reference to analyze inflationary pressures (which are linked to a positive value of output gap), and form an indicator to locate the economy in the business cycle.

2.2.1 Conceptual Issues

As of today, researchers still debate about the concept of potential output. First and foremost, there is not a clear and widely accepted definition of it: in a too narrow perspective, some refer to it as the highest level of GDP not causing inflationary pressures (non-inflationary output gap). Following a similar definition, some international institutions define it as the level of output corresponding to the NAIRU (Non-accelerating inflation rate of unemployment); another concept is that of sustainability, with researchers referring to the level of output sustainable by the economic structure. However, being estimated using economic modeling theory, this view is clearly too model-dependent. Even though standard macro models recently tend to agree on defining potential output as that corresponding to full employment, or when prices are fully flexible, the debate is still open and there is no agreement nor a precise and shared definition of potential output.

Second, its estimation presents different problems, and researchers do not agree on how to measure it. Not even the IMF or the ECB have an official method to compute it. Its estimation is usually computed by smoothing out business cycle fluctuations by measuring the trend component in the actual GDP series. However, even if there exist reliable statistical tools to isolate the cycle, estimating potential output in real-time brings poor results, since the trend can be estimated only relying on past data.

Clearly, neither rough information about real-time potential output nor precise information about past estimates are useful, so researchers debate on how to increase the accuracy of real-time measurements.

Potential GDP is strongly data-dependent, and very sensitive to the inclusion of new data. There is little doubt that in the years before the recent crisis many economies were running above potential (US surely was, after ex-post potential output revisions, Gavin, 2012), but the level of inflation remained low for a prolonged period of time (which is also one of the reasons why the non-inflationary view of potential output is addressed as too narrow). When dealing with such an imprecise and not robust estimate there seems to be an underlying problem, either in the statistical method used or in the validity of the relation estimated. The conceptual relationship between potential output and inflation implied by the Phillips curve is strong, despite the recent flattening.[1] So if the estimates of potential output are as we argued (ECB, 2009), then the problem must be in the method used to measure it.

2.2.2 Methodological Issues

Even though in the last few years the research on the estimation methodologies has partly moved from the academic community to international institutions, there is still not a widely accepted way of calculating potential GDP.

Researchers in the literature adopt two different approaches to estimate po-

[1]The higher output gap, the stronger inflationary pressure on prices, and vice-versa.

tential GDP: on the one hand there are univariate statistical approaches, which usually consist of filtering out the trend component from the cyclical one; on the other there are the structural approaches, which derive the estimates directly from the theoretical structure of a model.[2] The most widely used univariate statistical approaches are the H-P filter (Hodrick Prescott, 1997[3]), the Beveridge-Nelson decomposition,[4] and the Baxter-King filter.[5]

Despite being simpler than structural approaches, these methods all suffer from the so called end-point problem: they are extremely sensitive to the addition of new data and to real-time data revisions.

The alternative to univariate statistical methods are structural approaches, which are based on the implications and restrictions of the models built; these measures do not suffer from the end -point problem, but vary greatly across models.

It is also common to find hybrid approaches, mixing the two perspectives together. For example the OECD uses a method which stands somewhere in between a univariate approach (mainly the H-P Filter) and a model-built measure (relying on some economic relations to calculate for example the NAIRU[6]) to calculate trend participation rates, trend hours worked and trend total factor productivity

[2]We could also identify simple trend models, which assume the trend component to follow a particular function of time, for instance linear in logs as

$$ln(Y_t) = \alpha + \beta t + \epsilon_t \tag{2.1}$$

[3]This method estimates an unobserved variable and consists of solving a simple minimization of an objective function of the form

$$Min \sum_{t=1}^{T} (y_t - y_t^*)^2 + \lambda \sum_{t=2}^{T-1} [(y_{t+1}^* - y_t^*) - (y_t^* - y_{t-1}^2)]^2 \tag{2.2}$$

Implying a trade-off between goodness of fit to the actual series (first term) and degree of smoothness of the trend series (second term).

[4]Which extracts cycle and trend from a series by imposing restrictions: cycle and trend are assumed to be negatively correlated, and trend is assumed to follow a random walk path.

[5]Which in order to isolate the business cycle removes from the data both low and high frequencies: it filters out irregular high-frequency components and slow moving trend components.

[6]Non-Accelerating Inflation Rate of Unemployment, referring to some sort of modernization of the concept of Natural Level of Unemployment, the level of unemployment below which inflations pressures arise.

(Cotis et al., 2004).[7][8]

Around ten years ago the European Union changed its method of computing potential output to one very close to the OECDs (Denis et al., 2002) using a Cobb-Douglas production function with an exogenous trend.

The International Monetary Fund does not have an official method for computing potential output, and every country desk decides which measure fits best. The most common IMF approach uses a production function approach, with assumptions that vary greatly across countries, but discretion is left to the country desks which can decide otherwise.[9]

The different approaches of supra-national institutions can be very different, but generally share a common basis: the concept of a macroeconomic production function, which splits the contributions to potential output between changes in the use of the key inputs of capital and labor in the economy, and changes in their productivity what is usually called Total Factor Productivity (TFP). This last measure is a good indicator of technological progress and innovation of the economy.

2.3 A New Way of Thinking Potential Output

If one analyzes how many factors and determinants contribute to build the potential output of an economy, it appears that the statistical measures developed up to now (and in particular the univariate ones) are too simplistic and unsatisfactory.

The evolution of output gap depends directly on supply conditions: in the short run potential fluctuates because of variations of the endowment of the economy, the key inputs of capital and labor, and of their productivity and degree of utilization. These factors are, in turn, directly linked to labor market trends, technology innovation, and variations in the degree of investment. In the long run, the capacity of an economy is shaped by its legal, institutional, and economic framework: not only its set of laws, its tax system, its market regulations, the

[7]This method, however, is highly correlated to an output gap estimated by using the H -P filter (correlation close to 0.9 for all the G7 countries, apart from Germany for which it is 0.4.

[8]This method is explained in details in Giorno et al., 1995.

[9]For instance among the methods chosen for the United States we find the HP filter, the split time trend, the band pass filter.

stability of its monetary system and the efficiency of its responses, but also the quality of its educational system, demographic factors, financial market trends and many other features. Thus, in the long run potential output depends on the level of development of the country.

In the most simplistic way, potential and actual output differ when potential and actual production differ. In the literature, the benchmark for the sustainability of the path of output is inflation; however, this view is too narrow since inflation can remain low for a prolonged period of time, even though output is following an unsustainable growth path (as the recent economic crisis has shown).

Output gap is not measureable directly, and the contribution of all these factors cannot be quantified with certainty. Moreover, potential output depends on a large number of different factors, which contribute in different directions and with different magnitudes to modify its path. The reasoning that wider ideas of potential output and output gap are desirable, especially those embedding information about the financial cycle, is thus straightforward.

In a recent paper, Borio et al. (2013) developed an innovative way to embed more information in the estimation process for potential output. We quickly summarize their methodology:

They start with a state space model of the form:

$$\Delta y_t^* = \Delta y_{t-1}^* + \epsilon_{0,t} \tag{2.3}$$

$$y_t = y_t^* + \epsilon_{1,t} \tag{2.4}$$

In this simple representation, which specifies output gap for the standard HP filter, the difference between actual and potential output is just a normally distributed error term.

The basic model can be rewritten in companion form as:

$$\begin{pmatrix} y_t^* \\ y_{t-1}^* \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_{t-1}^* \\ y_{t-2}^* \end{pmatrix} + \begin{pmatrix} \epsilon_{0,t} \\ 0 \end{pmatrix} \tag{2.5}$$

$$y_t = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} y_t^* \\ y_{t-1}^* \end{pmatrix} + \epsilon_{1,t} \tag{2.6}$$

Where

$$Q = E\left[\begin{pmatrix} \epsilon_{0,t} \\ 0 \end{pmatrix} \begin{pmatrix} \epsilon_{0,t} & 0 \end{pmatrix}\right] = \begin{pmatrix} \sigma_0^2 & 0 \\ 0 & 0 \end{pmatrix} \tag{2.7}$$

$$R = E\left[\epsilon_{1,t}\epsilon_{1,t}\right] = \sigma_1^2 \tag{2.8}$$

where $y_t = ln(Y_t)$ is the natural logarithm of real GDP, and $\epsilon_{0,t}$ and $\epsilon_{1,t}$ are assumed to be two normally and independently distributed error terms with mean zero and unknown variances σ_0^2 and σ_1^2 , respectively. The ratio between the variances of the error terms in the state and measurement equation is nothing but the lambda term of the HP filter (the so called noise-to-signal-ratio) $\lambda = \sigma_1^2/\sigma_0^2$, which is set to 1600, as usual for quarterly data.

The idea of Borio, Disyatat and Juselius to enrich this estimate, is to apply the Hodrick-Prescott filter by means of a Kalman filter, including in the model some proxies for the financial cycle.

In order to include information about the financial cycle in the estimates of the output gap, the authors augment equation (2.4) with a vector including additional variables:

$$y_t = y_t^* + \gamma' x_t + \epsilon_{2,t} \tag{2.9}$$

Where x_t is the vector containing these additional variables: lags of the output gap itself, indicators of house prices, interest rates, credit, while $\epsilon_{2,t}$ is a normally and independently distributed error term with mean zero and unknown variance σ_2^2.

Now, using an extended version of the measurement equation (2.9), in order to keep the same duration of the business cycle as the standard HP filter the authors use the state equation (2.3), and set the new noise-to-signal ratio $\lambda_2 = \sigma_2^2/\sigma_0^2$ such that:

$$\frac{Var(y_t - y^*_{(2),t})}{Var(\Delta^2 y_{(2),t})} = \frac{Var(y_t - y^*_{(3),t})}{Var(\Delta^2 y_{(3),t})} \tag{2.10}$$

Where $y_{(2),t}$ and $y_{(3),t}$ are the output gap series obtained from the measurement equations (2.4) and (2.9), respectively.

The authors consider many different specifications of (2.9), including different variables in the vector x_t. In its final version, the measurement equation has the form:

$$y_t - y^*_t = \beta(y_{t-1} - y^*_{t-1}) + \gamma_1 r_{t-k_r} + \gamma_2 \Delta cr_{t-kcr} + \gamma_3 \Delta ph_{t-kph} + \epsilon_{3,t} \tag{2.11}$$

Where $r_t = i_t - \Delta p_t$ is the ex post real interest rate, Δp_t is the first difference of the natural logarithm of the consumer price index, Δcr_t is the percent growth of real credit to the non-financial private sector, Δph_t is the percent growth of the real residential property price index, and $(y_{t-1} - y^*_{t-1})$ is an autoregressive component for output gap.

Each of these variables enters only once in (2.11) with a lag k_j, with $j = 0, 1, 2, 3, 4$ chosen to maximize the statistical fit. To clarify this point, the authors try one by one the first four lags of each variable, and choose to include in the final expression of the equation only the one which fits best.[10]

In order to estimate equation (2.11), Borio et al. adopt a Bayesian approach. The procedure employed is two-fold: first the unknown variances σ_0^2 and σ_3^2 are estimated via maximum likelihood; second, the Kalman filter is used in order to obtain the smoothed values of the unknown series y^*_t, using the maximum likelihood estimates of the variances $\hat{\sigma}_0^2$ and $\hat{\sigma}_3^2$. There is of course a loss of information here: by considering the sigmas as known values you do not consider, for instance, the variability of the estimates.

The authors, without any clear justification, assume the priors for all parameters to be gamma distributions with standard deviation equal to 0.2. They then also force β to lie between 0 and 0.95, with a prior mean of 0.80, while γ_i with $i = 1, 2, 3$ are restricted to lie between 0 and 1, with a prior mean equal to 0.2.

[10]For instance for the case of United States real interest rate is taken at the second lag, real credit with no lag and house price at the fourth lag: the ones that maximize the statistical fit.

They consider four different specifications of the simple model; the first four are represented in Table (2.3): a simple autoregressive one considering only potential output at the previous period, one considering the autoregressive component and interest rate, one including the autoregressive component and credit, and one including the autoregressive component and property price.

Looking at Table (2.3), the introductory results confirm that the proxies for the financial cycle are significant and contain important information about the business cycle. In particular, the coefficients on credit growth and property price are strongly significant in all the nations examined. Real interest rate is significant only in the case of United States, while the autoregressive component of the output gap is always highly significant. This result strongly contradicts that of equation (2.4), which represents the static HP filter, where output gap was simply represented by the error term.

The authors then turn to a more general analysis, where all the variables are included simultaneously in the model (but still only at the lag which maximizes the statistical fit), and non-linear relations are allowed.

Once again, the autoregressive output gap component is highly significant for all specifications and countries, while the interest rate is never significant: this finding confirms the results of the basic model. Both changes in property prices and credit growth are significant in the general model, sustaining the hypothesis of the authors that proxies for the financial cycle carry information about the cycle.

Interestingly, the new estimates show much larger deviations of output from sustainable levels during the early 2000s, the period before the economic crisis. Since that was a time of increasing private sector leverage, the authors draw the implicit conclusion that their new estimation method performed better during the last decade. Even though the causal relationship might be called into question, estimates considering a wider concept of output gap are more precise and perform better than simple univariate statistical methods.

The coefficients for real interest rates are significantly negative in all the countries examined. This result suggests that the higher the ex-post real interest rate, the lower the output gap will be. The coefficients for the autoregressive component, credit and house price index are instead positive. But more than signs and magnitudes of the coefficients, the key here is to underline that financial factors

Table 2.1: Regression results: individual explanatory variables

	United States				United Kingdom				Spain			
Model	1	2	3	4	1	2	3	4	1	2	3	4
β	0.95	0.90	0.82	0.91	0.95	0.95	0.94	0.88	0.95	0.95	0.90	0.95
	(-)	(14.33)	(14.74)	(17.0)	(-)	(-)	(15.96)	(12.89)	(-)	(-)	(14.57)	(14.35)
r	-	-0.08	-	-	-	-0.02	-	-	-	-0.03	-	-
		(3.79)				(-0.85)				(-1.85)		
Δ cr	-	-	0.58	-	-	-	0.10	-	-	-	0.15	-
			(6.30)				(3.73)				(2.99)	
Δph	-	-	-	0.17	-	-	-	0.11	-	-	-	0.07
				(5.48)				(4.15)				(2.87)
kr	-	-2	-	-	-	-1	-	-	-	0	-	-
kcr	-	-	0	-	-	-	0	-	-	-	-2	-
kph	-	-	-	-4	-	-	-	-2	-	-	-	-3

Table 2.2: Regression results: full specifications

	United States			United Kingdom			Spain		
Model	1	2	3	1	2	3	1	2	3
β	0.81	0.80	0.81	0.88	0.87	0.88	0.90	0.89	0.84
	(14.13)	(14.30)	(15.62)	(14.03)	(14.70)	(15.28)	(10.88)	(15.00)	(15.47)
r	-0.04	-	-	-0.03	-	-	-0.03	-	-
	(5.06)			(-1.20)			(-1.85)		
Δcr	0.51	0.52	0.62	0.09	0.09	0.12	0.12	0.13	0.54
	(5.06)	(5.56)	(5.59)	(3.54)	(3.81)	(3.79)	(2.25)	(2.91)	(4.85)
Δph	0.09	0.10	0.12	0.10	0.11	0.11	0.06	0.06	0.03
	(2.69)	(2.78)	(2.34)	(3.79)	(4.29)	(3.61)	(2.43)	(2.51)	(2.59)
kr	-2	-	-	-1	-	-	0	-	-
kcr	0	0	0	0	0	0	-2	-2	-2
kph	-4	-4	-4	-2	-2	-2	-3	-3	-3
τcr	-	-	0.044	-	-	0.024	-	-	0.017
τph	-	-	0.024	-	-	0.019	-	-	0.015
ρcr	-	-	9.85	-	-	18.50	-	-	31.90
ρph	-	-	17.60	-	-	28.35	-	-	42.35

carry information to explain potential output, and that embedding such information leads to better estimates.

The new estimates are in fact more precise than the standard H-P filter ones (equation (2.4)): by comparing 95% confidence intervals for the output gap estimates, the sizes of these intervals are significantly smaller with the new estimation method.[11]

The finance neutral estimates[12] are also less subject to ex-post revision: Borios measure of output gap follows the ex-post gap much more precisely than other estimates which do not take into account financial factors. Moreover, the new estimation method produces a series which is remarkably more sensitive in detecting unsustainable booms like the one which preceded the 2008 global economic crisis.

The last four rows of the table represent non-linearities. Financial variables are weighted in equation (2.11) according to the size of the underlying imbalances (estimated relying on Borio and Drehmann (2009)).

The results for non-linearities are significant, but not as much as the authors expected. They do not have a strong impact on any of the countries analyzed.

2.3.1 A wider jurisdictional scope

As a fist step into the extension of such model, we here widen the initial analysis which considered only three countries (United States, United Kingdom and Spain), to include more nations and a longer time series.

We find similar results by replicating the analysis with the new dataset.[13]

Following the original methodology, we consider logarithmic first differences of real variables, assign the same gammas as prior distributions to the parameters, and solve the model. First we consider the property price neutral model, in which we include real property price in the estimation of potential output, then the credit neutral version of the model in which we include the logarithm of real credit to the private sector, and finally the finance neutral one, fully specified with both of

[11] More than halved for United States, form ± 3.50 to ± 1.35, and notably smaller both for Spain (from ± 3.85 to ± 2.10) and United Kingdom (from ± 2.95 to ± 1.80).

[12] Those in which proxies for the financial cycle are included.

[13] We replicate the original paper for Australia, Austria, Canada, France, Greece, Italy, Japan, Spain, Switzerland, United Kingdom and United States.

Table 2.3: Regression results: full specifications

	United States		United Kingdom		Spain	
Model	1	2	1	2	1	2
β	0.88	0.95	0.92	0.93	0.76	0.95
	(16.31)	(26.21)	(15.00)	(19.95)	(7.92)	(4.88)
Δ cr	0.32	-	0.04	-	0.14	-
	(3.69)		(0.90)		(4.40)	
Δph		0.22		0.14		0.11
		(4.50)		(4.60)		(2.42)
kcr	0	-	0	-	-4	-
kph	-	-3	-	-1	-	-1

Table 2.4: Regression results: individual explanatory variables

	UK		Spain		Italy		France	
Model	1	2	1	2	1	2	1	2
β	0.92	0.93	0.76	0.95	0.92	0.93	0.95	0.95
	(15.00)	(19.95)	(7.92)	(4.88)	(15.59)	(19.68)	(22.88)	(11.53)
Δ cr	0.04	-	0.14	-	0.01	-	0.01	-
	(0.90)		(4.40)		(0.35)		(0.32)	
Δph		0.14		0.11		0.30		0.06
		(4.60)		(2.42)		(4.96)		(2.21)
kcr	0	-	-4	-	-3	-	0	-
kph	-	-1	-	-1	-	-3	-	0

the proxies for the financial cycle.

Our results show significant differences among countries. We confirm the results of the original paper for United States and Spain (even though we find that the statistical significance is somewhere maximized for different lags), while for United Kingdom we do not find any significance for real credit.

As extention to more jurisdictions, we performed the same calculations for Italy, France, Netherlands, Austria, Switzerland, Canada and Australia. The single explanatory variable estimations are presented in Tables 2.4, 2.5 and 2.6.

For the states where both the proxies are significant, we also perform the full-specification regression, reported in Table (2.6).

We confirm the results for the three countries initially considered by Borio et al., except for the case of United Kingdom, for which we do not find any statistical significance for credit.

Table 2.5: Regression results: individual explanatory variables

	Austria		Netherlands		Switzerland	
Model	1	2	1	2	1	2
β	0.50	0.29	0.72	0.23	0.95	0.95
	(5.94)	(3.78)	(7.13)	(3.17)	(1.97)	(5.20)
Δ cr	0.26	-	0.42	-	0.05	-
	(8.73)		(10.09)		(2.47)	
Δph		0.50		0.08		0.07
		(5.03)		(0.66)		(2.07)
kcr	0	-	0	-	0	-
kph	-	-4	-	-2	-	-2

Table 2.6: Regression results: individual explanatory variables

	US	Spain	Switzerland	Austria
β	0.87	0.78	0.95	0.47
	(16.01)	(7.02)	(0.24)	(5.83)
Δ cr	0.21	0.14	0.05	0.23
	(2.54)	(4.49)	(2.64)	(8.02)
Δph	0.20	0.09	0.07	0.31
	(3.75)	(2.15)	(2.39)	(3.20)
kcr	0	-4	0	0
kph	-3	-1	-2	-4

This result is quite consistent across countries: the property price index is highly significant in almost all the countries analyzed (apart from the Netherlands), while real credit seems to have a more limited role.

In Italy only real property price is significant, and the same result arises for France, Australia, and Canada. The Netherlands represent an exception: real property price does not carry useful information, while credit is highly significant.

One notable fact is that property price is usually most significant in lagged value, probably reflecting the lag with which the housing market responds to economic shocks.

These results suggest that the financial cycle proxies do not carry the same information in all countries, so we cannot draw common or general conclusions. Credit and real property price surely ameliorate the estimation of potential output, but the source and magnitude of this improvement varies across countries and thus has to be analyzed case by case.

The fact that some of our estimates differ from those of Borio et al. might arise from the fact that we used longer time series (see Appendix B for a precise description of the data used),[14] but could also mean that the theory underlying the idea of the original paper is not very robust: in fact we do not find the same results regarding the lags,[15] and significance of the variables, and do not find any statistical significance in some of the countries analyzed (Australia, Canada, Greece and Japan).

2.4 Extension: Short-term Debt

Credit and property price are just two of the possible proxies that could be included in the estimation procedure. The preceding analysis raises the question of whether other variables might carry significant information in estimating potential output, and thus might be used in the model specified in the previous paragraphs.

One potentially interesting variable to introduce into the model is a measure of liquidity. We strongly believe that liquidity plays a central role in informing about the health of the financial sector: it is important for analyzing both the level of systemic risk (aggregate measures of liquidity are helpful to detect systemic risks during a period of growth) and the financial situation of economic agents (from single agents to firms).

This concept becomes even more important when analyzing a period of financial crisis, since liquidity is a key response indicator: as Brunnermeier (2013) states, market participants have different reactions to shocks, depending on whether or not they face liquidity problems.

The aspect of liquidity we want to account for in the model is that of short-

[14] Apart for United Kingdom for which we used shorter time series, and in fact the results we find differ with respect to the reference paper.

[15] Once again, every specification of the model considers for credit and property price only the lag among the first four which maximizes the statistical fit.

term debt (normalised taken as a ratio over GDP). The trend of the most liquid liabilities before a crisis before a period of economic troubles when many assets will become illiquid seems to be a good indicator of contingent underlying tensions of the economy. This does not necessarily imply a cause-effect relation, in the sense that growth in the amounts of short-term debt should increase the vulnerability of the economy, but instead implies that short-term debt is a good indicator of the imbalances in the economy which might then result in a financial crisis. In this sense, our view is not in contrast with Benmelech and Dvir (2010), which see short term debt as reflecting, rather than causing distress in the banking sector, and is perfectly in line with Krishnamurthy and Vissing-Jorgensen (2012) when saying that short-term debt (in particular that issued by financial institutions) is a good predictor of financial crises.

Our final goal is to further filter out the financial cycle: according to the BIS, if policymakers make their decisions focusing only on the business cycle, some imbalances might arise, such as overindebtedness in the corporate or household sectors. Particularly in the private sector, high debt levels "can undermine sustainable economic growth" (BIS (2014)).

2.4.1 Liquidity estimates for United States

It was with surprise that we discovered an impressive lack of data on short-term debt. Our initial idea was to consider the amount of liabilities due within one year as a proxy for the financial cycle, and to analyze their contribution disaggregating data by sector of debt issuer. A sectorial analysis would in fact give a better understanding of the relationship between short-term debt and financial cycle. However, because of the lack of data on the proxy we have identified, our idea immediately revealed insurmountable limits: central banks hardly publish such data, and the available series for debt are not disaggregated, neither by maturity nor by sector.

The only country for which we have been able to retrieve a satisfactory series of data (i.e. long enough to allow us to perform a complete analysis and diagnostic) is United States. We take short term debt of the nonfarm and nonfinancial corporate business as a proxy for the financial cycle, which we include in equation (2.9). The

Table 2.7: Regression results: extension with short-term debt

	United States		
Model	1	2	3
β	0.78	0.81	0.82
	(12.73)	(15.33)	(15.30)
Δ nfstd	0.08	0.04	0.06
	(3.94)	(2.21)	(3.20)
Δ cr	-	0.23	-
		(3.14)	
Δ ph	-	-	0.22
			(4.29)
k nfstd	0	0	0
k cr	-	0	-
k ph	-	-	-3

results of this specification of the model are reported in Table 2.7.

As in the previous paragraphs, we estimate the model successively including short-term debt at different lags. The coefficient of the proxy we include is positive and highly significant: it seems that United States debt due within one year is a valid proxy for the financial cycle, confirming our initial hypothesis. Moreover, the statistical fit is maximized in the model with the unlagged series for debt.

We also perform the regression of short term debt with credit and property price, the proxies we used in the previous analysis, and still find statistical significance. Estimating the model with the augmented specifications, we confirm all the results we obtained analyzing one variable at a time: all variables still have strong significance, and the statistical fit is maximized at the same lags of the case with the individual explanatory variables.

As we can see from figures 2.1 to 2.3, our measure of output gap estimates the economy to be over potential before the crisis, flattening then in proximity of the recession as if it was anticipating some sort of imbalance of the economy. The behavior is different with respect to the simple HP filtered series, whose estimate of potential output keeps on growing during the crisis.

Another important feature of our estimation method is that it seems to be less subject to ex-post revision, and to the addition of new data than simple filtering.

Figure 2.1: United States debt neutral output gap with confidence intervals

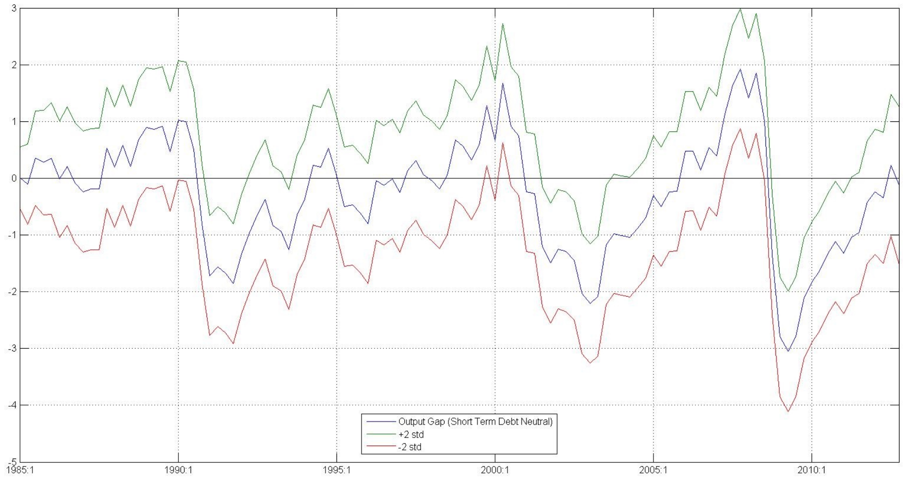
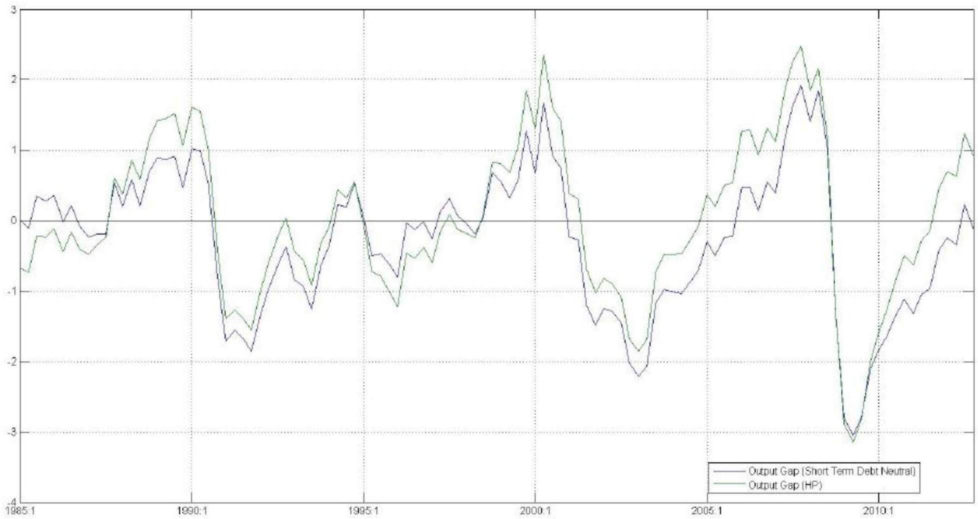


Figure 2.2: United States output gaps



Especially around the end of the sample, the "debt neutral" measure of output gap seems to perform better, when the sample is extended to include new data. Figures 2.4 and 2.5 show the differences in ex-post revisions between the simple HP-filtered series and the debt neutral one.

Figure 2.3: United States actual and potential outputs

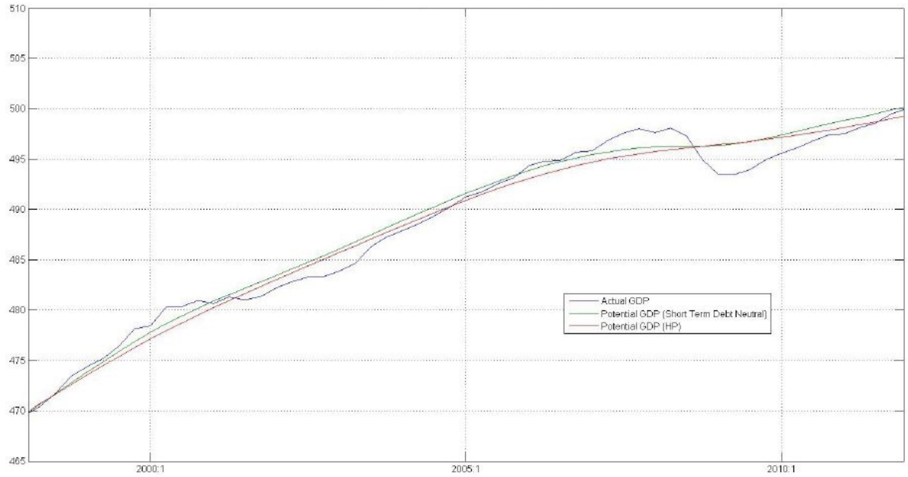
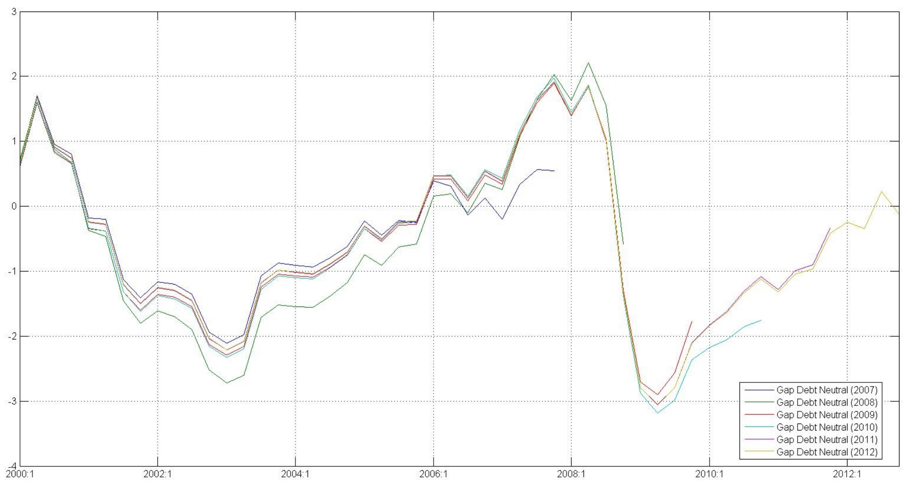
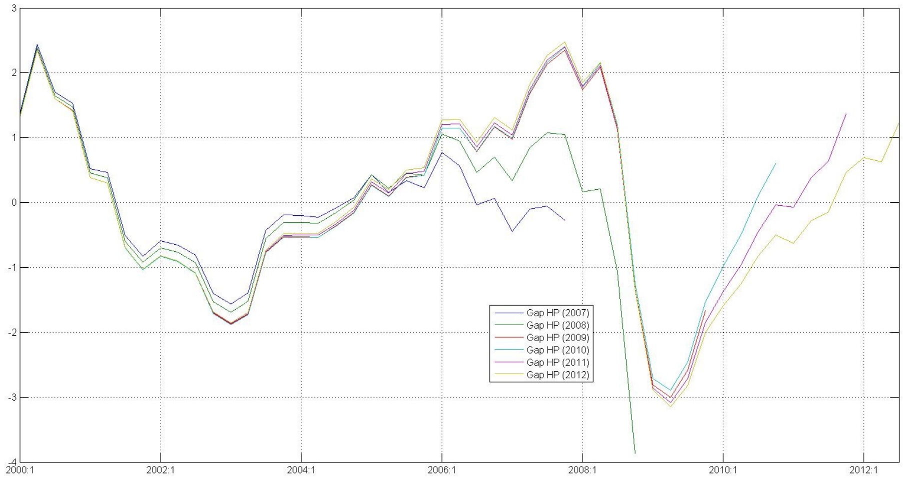


Figure 2.4: United States debt neutral ex-post output gap



Note from the figures how in the case of the HP filtered output gap the robustness of the estimates is much lower. Especially at the end of the sample, the extended model including short-term debt performs better and is subject to smaller revisions.

Figure 2.5: United States HP filtered ex-post output gap



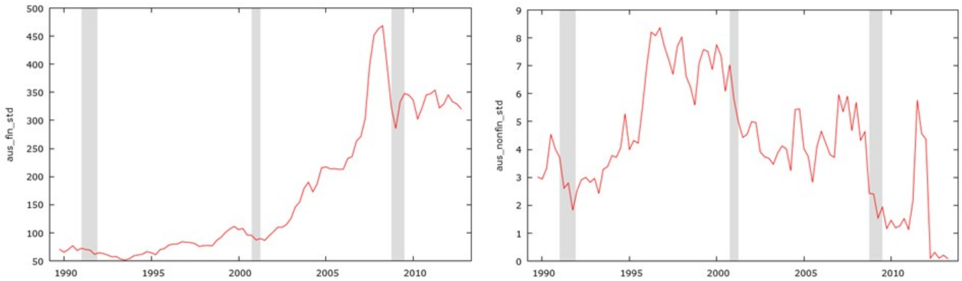
2.4.2 Analysis for other countries

Due to a lack of data we have been unable to construct satisfactory data series for short-term debt for many countries.

Apart from United States, for which government data are freely available, we use the BIS dataset on short-term amounts outstanding by sector. These series depict the path of those liabilities whose original maturity was within 12 months, and are disaggregated by sector according to three macro-categories: general government, financial and nonfinancial corporations. The main two problems with these series are the length (the longest ones start in the fourth quarter of 1989) and incompleteness (data for many developed countries are not collected, and for others the series is made of only a handful of observations). For these reasons we limit our analysis to the two countries for which data are satisfactory and permit at least a good preliminary analysis: Australia and Canada.

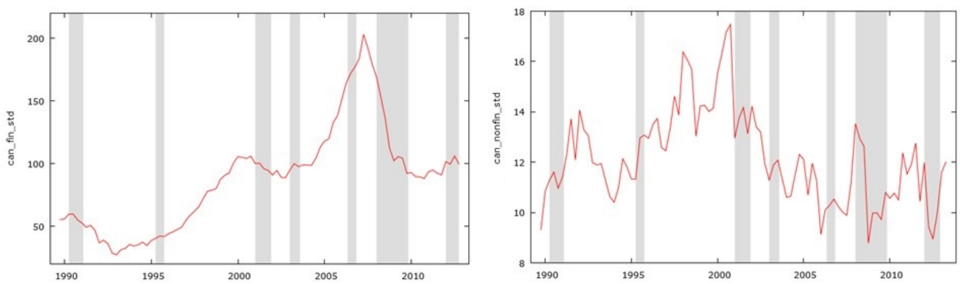
We start by drawing some quick stylized facts from the graphical representations of the BIS series, in order to justify our work (the graphs on the left represent short-term debt for the financial sector, while the ones on the right debt for the non-financial sector; grey bars represent recessions or periods of almost zero GDP growth).

Figure 2.6: Australia short-term debt for financial and non-financial sectors. Source: BIS



The graphs for Australia show three periods in which the economy has slowed down, and (especially for the non-financial sector) there has been a clear peak of debt before every GDP slowdown.

Figure 2.7: Canada short-term debt for financial and non-financial sectors. Source: BIS



Even in the case of Canada many quarters of GDP decrease have been preceded by a peak of short-term debt. In particular regarding the recent 2008 global crisis, the peak-drop pattern of debt in both financial and non-financial sector is evident. But despite being relatively smaller in absolute numbers, a similar pattern can be identified even in previous periods, especially for the financial sector .

The case of Japan is interesting as well, but we analyze it only graphically since the series for short-term debt is too short: it consists only of a bunch of observations, and does not permit a satisfactory statistical analysis. The time series is missing almost the entire 1990s decade, a period which would be particularly interesting to analyze. In any case the data we have seem encouraging, even though these short series make it difficult to perform satisfactory estimates.

Figure 2.8: Japan short-term debt for financial and non-financial sectors. Source: BIS

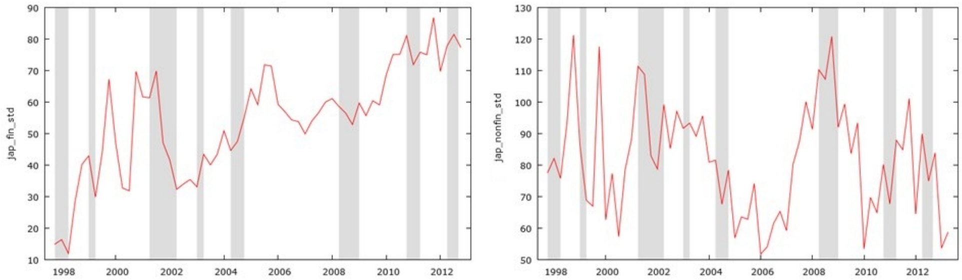


Table 2.8: Regression results: extension with short-term debt

	Australia		Canada	
Model	1	2	1	2
β	0.57	0.64	0.88	0.85
	(4.59)	(5.00)	(13.40)	(12.58)
Δ fstd	0.02	-	0.02	0.02
	(2.09)		(2.01)	(2.02)
Δ nfstd	-	0.004	-	
		(1.53)		
k fstd	2	-	2	3
k nfstd	-	0	-	-

In any case, the peak-drop pattern is once again clearly evident.

Analyzing the case of Australia and Canada in Table (2.8), we can get a more general picture of the role that short-term debt plays in our output gap estimates.

For both of the countries we find slight significance of short-term debt of the financial sector. We instead find statistical significance for the nonfinancial sector only in the case of United States.

Recall that for both Canada and Australia we have not found any statistical significance for the financial cycle proxies (credit and property price), but the results for short term debt (of the financial sector) are encouraging, despite the slight significance.

These two countries represent a particularly interesting case, being advanced economies that were not particularly hit by the recent global economic downturn despite having had large financial booms in the late 2000s. This process of financial expansion has just slowed down because of the global crisis, and the fact that households continued to borrow (even though at a slower pace) together with the strong increases in commodity prices prevented a lasting turn of the cycle, causing real property prices to be back to the high (and possibly overrated) levels of the years of the boom.

As already pointed out, according to BIS data, countries are at very different stages of the financial cycle. This pattern is consistent with our findings that output gap measures must be personalized by country, and there is not a common and shared dogmatic measure that works for every nation. This is also indirectly in line with the approach of the IMF, where discretion is left at every country desk to choose the estimation method that fits best.

Even though liquidity conditions are often highly correlated across countries, the financial cycle seems to take a different shape in every country. And it is by filtering out the proxies identified (sometimes credit to GDP and property price, sometimes debt to GDP) that we make a first step in retreiving a reliable measure of "finance neutral" potential output.

The key statistical problem of these last estimates is the length of the BIS series used to compute them. Both Borios work and our previous replications rely on quarterly data of at least 35 years, while the series available for short-term debt (apart for United States) date back to 1989q4 for Australia and Canada, and 1997q4 for Japan. Working with time series data, the span is far more important than the density of data, and we cannot say that our dataset is satisfactory from this point of view.

However, these preliminary results are encouraging, and we indeed find that short-term debt carries information regarding the financial cycle. Unfortunately, it is not easy to construct short-term debt time series for every nation: not all states collect quarterly data for short-term privately held liabilities. Moreover, even where available, these measures might be extremely heterogeneous among states, this complicating inter-state comparisons.

2.5 Conclusion

As we have seen, the concept of output gap has evolved much over time. The raw definitions of the 1960s have now developed into precise statistical measures, on which central banks and governments rely to take policy measures. However, we have argued that, despite being theoretically and statistically rigorous, todays measures to estimate potential output and output gap are not satisfactory, and wider estimation methods are needed.

On this behalf, we have started from the work of Borio et al. (2013), which includes proxies for the financial cycle to estimate potential output, analyzing and replicating their results in order to check the robustness of this innovative method. By replicating their analysis for more countries, we have found that financial information is important in explaining the potential of the economy, even though there is not a unique and precise rule to describe its role across countries.

All of our results (together with the most recent literature on the topic) suggest that the Phillips curve represents a less strong relation than it used to. As a consequence, if inflation has become less responsive to variations in output, and the Phillips Curve is in crisis, then new methods to estimate the output gap are needed.

The new estimates presented in this work suggest that finance neutral and "debt neutral" estimates perform better than traditional methods (H-P filter or production function approaches): including proxies for the financial cycle is a good way to improve the precision of the estimates of output gap.

We have also shown that credit and property price are not the only proxies for the financial cycle that perform well in this framework: even short-term debt carries statistical significance in estimating output gap via Borios innovative method.

The bottom line is that even though the general estimation method is applicable to different countries, the result shaped by the data is diverse across nations. There is not a common and univocal measure that works for every country.

There are, however, many issues that are still open: despite showing strong statistical significance, the choice of the lags in the estimation procedure is quite arbitrary, and there is no justification apart from that of simply choosing the lag that maximizes the statistical fit. According to our computations, variables show statistical significance even at higher lags, in particular for housing prices (e.g. residential property price is highly significant for the US at lag 12, once again reflecting the lag with which the housing market responds to economic shocks).

Moreover, the estimates come from a simple state-space model, and not from a fully specified macroeconomic model, this preventing us from fully interpreting the result and conducting deeper analyses.

Further developments might also include testing for breaks and regime switches in the series and their eventual consequences on the estimates, the study of time-varying coefficients, and the development of a structural microfundeed model with financial frictions.

2.6 Appendix A: Variables specification

For sake of clarity, we specify *in extensor* all the variables and parameters included in the tables.

β is the coefficient for the AR component for output gap $(y_{t-1} - y^*_{t-1})$,

r represents the coefficient for the ex post real interest rate $r_t = i_t - \Delta p_t$,

Δp_t is the first difference of the natural logarithm of consumer price index,

Δcr_t is the % growth of the real credit to the non-financial private sector,

Δph_t is the percent growth of the residential property price index,

$\Delta nfstd_t$ is short-term debt of the non-financial sector,

$\Delta fstd_t$ is short-term debt of the financial sector,

k_r represents the lag considered for the ex post real interest rate,

$k_c r$ is the lag for credit to the non-financial private sector,

$k_p h$ represents the lag for the residential property price index,

$k_n fstd$ represents the lag for short-term debt of the non-financial sector,

$k_f std$ represents the lag for short-term debt of the financial sector,

ρ and τ represent the coefficients for non-linearity.

Moreover, since many variables show some cyclicality, the authors calculate their averages by Cesro means: this is done by building a mean sequence in which the sample is successively increased by one observation. By doing so, the convergence is much faster and the pro-cyclicality in the mean-adjustment is reduced.

2.7 Appendix B: Data specification

The time series we have used were based on the following data:

GDP_t: seasonally adjusted nominal gross domestic product in own currency. Source: OECD Economic Outlook.

$PGDP_t$: GDP-deflator. Source: OECD, Main Economic Indicators.

RPP_t: Residential property price. Sources: Australian Bureau of Statistics for Australia, OECD data (Main Economic Indicators and Economic Outlook) for Austria, Netherlands, UK and Italy, Oxford Economics for the remaining countries.

CR_t: Credit to private non-financial sector. Source: BIS Data.

CPI_t: Consumer price index. Source: OECD, Main Economic Indicators.

STD_t: Short term debt. Sources: for United States Federal Reserve Short-term debt of Nonfarm Nonfinancial Corporate Business. For other nations BIS data.

Bibliography

[1] Basu S. and Fernald J.G., 2009, *What Do We Know (And Not Know) About Potential Output*, Federal Reserve Bank of St. Louis, *Review*, 91(4), 187-213.

[2] Benmelech E. and Dvir E., 2010, *Does Short-term Debt Increase Vulnerability to Crisis? Evidence from the East Asian Financial Crisis*, Journal of International Economics, 89 (2013), 485-494.

[3] BIS, 2014, *84th BIS Annual Report, 2013/2014*, Bank for International Settlements.

[4] Borio C., Drehmann M., 2009, *Towards an Operational Framework for Financial Stability: "Fuzzy" Measurement and its Consequences*, BIS Working Papers, n.284.

[5] Borio C., Tsatsaronis K., 2012, *Characterising the financial cycle: don't lose sight of the medium term!*, BIS Working Papers, n.380.

[6] Borio C., Disyatat P., Juselius M., 2013, *Rethinking Potential Output: Embedding Information about the Financial Cycle*, BIS Working Papers, n.404.

[7] Boissay F., Collard F. and Smets F., 2013, *Booms and Banking Crises*, ECB Working Paper Series n.1514.

[8] Brunnermeier M., 2008, *Deciphering the Liquidity and Credit Crunch 2007-2008*, NBER working papers 14612, National Bureau of Economic Research, Inc.

[9] Brunnermeier M., Krishnamurthy A. and Gorton G., 2013, *Liquidity Mismatch Measurement*, NBER Chapters, in: Risk Topography: Systemic Risk and Macro Modeling, National Bureau of Economic Research, Inc.

[10] Cahn C. and Saint-Guilhem A., 2007, *Potential Output Growth in Several Industrialized Countries, a Comparison*, ECB Working Paper Series n.828.

[11] Congdon T., 2008, *Two Concepts of Output Gap*, World Economics, 9(1), 147-75.

[12] Cotis J.P., Elmeskov J. and Mourougane A., 2004, *Estimates of Potential Output Benefits and Pitfalls from a Policy Perspective*, in Reichlin L., *The Euro Area Business Cycle: Stylized Facts and Measurement Issues*, CEPR, London.

[13] Denis C., Mc Morrow K. and Roger W., 2002, *Production Function Approach to Calculating Potential Growth and Output Gaps Estimates for the EU Member States and the US*, ECFIN Economic Paper, n.176.

[14] European Central Bank, 2009, ECB Monthly Bulletin, July 2009.

[15] European Central Bank, 2011, *Trends in Potential Output*, ECB Monthly Bulletin, January 2011.

[16] Friedman M., 1968, *The Role of Monetary Policy*, American Economic Review, 58(1), 1-17.

[17] Gavin T., 2012, *What is Potential GDP and Why Does it Matter?*, Federal Reserve of St. Loius Economic Synopses, n.11.

[18] Giorno C., Richardson P., Roseveare D., and Van den Noord P., 1995, *Estimating Potential Output, Output Gaps and Structural Budget Balances*, OECD Economic Department Working Paper n.157.

[19] Haltmaier J., 2012, *Do Recessions Affect Potential Output?*, Board of Governors of the Federal Reserve System, International Finance Discussion Papers, n.1066.

[20] International Monetary Fund, 2013, *The Dog That Didnt Bark: Has Inflation Been Muzzled or Was It Just Sleeping?*, World Economic Outlook, Chapter 3, 79-95.

[21] Krishnamurthy A. and Vissing-Jorgensen A., 2012, *Short-term Debt and Financial Crises: What Can We Learn from U.S. Treasury Supply?*, Working Paper.

[22] Mayer L. H., 2001, *Does Money Matter?*, Federal Reserve Bank of St. Louis, 1-16.

[23] Montoya L. A. and Dohring B., 2011, *The Improbable Renaissance of the Phillips Curve: the Crisis and Euro Area Inflation Dynamics*, European Commission Economic Papers, n.446.

[24] Okun A. M., 1962, *Potential Gnp: Its Measurement and Significance*, Proceedings of the American Statistical Association.

[25] Orphanides A., 2003, *The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time*, Journal of Money, Credit and Banking, 37(3), 583-601.

[26] Phelps E. S., 1967, *Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time*, Economica, 34, 254-81.

Chapter 3

A New Estimation Method for Employment Trend

3.1 Introduction

Estimating the potential and cyclical component of labor market variables is of crucial importance for economic research. Recent studies have tried to quantify the effectiveness of policy intervention on the process of job creation (Monacelli et al. 2010, Bruckner and Pappa 2012, and Turrini 2013), and have focused on the ability of discretionary policies (both fiscal and monetary) to reduce the cyclical component of unemployment. This is usually estimated as the deviation of the unemployment rate from the NAIRU.[1] However, methodologies employed to identify the cyclical component, such as purely statistical filtering techniques (e.g. univariate or multivariate HP filters) and more structural model approaches (see Borio et al. 2014 for a discussion), suffer of some shortcomings, both conceptual and methodological.[2]

Regarding conceptual issues, simple filtering techniques struggle to capture the demographic and social factors that are strongly modifying the composition of the population, and will deeply change its future structure. Social phenomena such as population aging or the increase of female participation strongly affect

[1]Non-Accelerating Inflation Rate of Unemployment.

[2]This work is coauthored with Andrea Tafuro, Ca' Foscari University of Venice

the labor force; these are long-term dynamics that slowly modify and smooth the age composition of the labor force and push it towards a different equilibrium. The aggregate level of employment represents the average behavior of different population cohorts dynamics, which can be similar as well as divergent. Traditional estimation methods are unable to account for these different dynamics, and for their consequences on the labor market.[3]

On the methodological side, recent contributions by Borio et al. (2013 and 2014) underlined how augmenting the measurement equation with financial variables in a state-space framework to estimate potential variables (in their case GDP) potentially produces a more precise and robust estimate. However, as recent theoretical studies (Monacelli et al. 2011; Petrosky-Nadeau 2014; Garn 2015; and Miao et al. 2016) highlight, changes in credit conditions also directly affect the labor market. According to these studies, changes in credit conditions affect the bargaining power of firms: during credit booms firms are less constrained, this enlarging their bargaining power. Since firms willingness to hire increases when their bargaining power is high, the labor demand will be larger during credit booms. Therefore, we add to our state-space model some proxies for the financial cycle in order to filter out their possible contribution to cyclical employment movements.

What we propose here is a two-step estimation method for the employment trend that tries to solve some of these issues. We estimate trend employment using a Kalman filter procedure in a state-space framework, and conduct the estimation separately for each age-gender cohort. We then aggregate the gender-cohort specific series, to obtain the final time series for the population trend employment. This procedure allows us to account for the structural demographic changes society is currently experiencing.

The second important innovation lies in the state-space model formulation. We augment the measurement equation to include some proxies for the financial cycle, á la Borio (2012). In this way, we retrieve a cleaner and cycle-free estimated series that is significantly more precise and robust over time with respect to other methodologies commonly used - such as simple HP or Kalman filters applied on

[3]On this see e.g. Gordon (2014) and Hall (2014), whose works investigate how population aging can modify the labor market participation rates, and thus induce a different employment potential level.

employment as a whole: it reduces the indeterminacy of estimation, and produces a more accurate quantification of the cyclical and trend components (for instance, reducing the well-known end-point problem).[4]

A simple Kalman Filter (KF) that directly filters the total value of employment usually estimates a smaller fall in employment cyclical component during the Great Recession, with respect to the level we obtain aggregating the single cohorts estimates. This result confirms that a filter applied without considering the cohorts is unable to cope with structural changes in labor force due to demographic trends, and ends up underestimating the cyclical component when the share of elders is increasing over time.

Our methodology builds on Borio et al. (2012), which uses a similar technique to improve the estimation process of potential output. Their augmented KF embeds proxies for the cycle, and filters out eventual cyclcal components from the estimated series. We also include in our state-space specification some extra variables to filter out the cyclical component of the employment trend, and we expect these variables to affect mostly the marginal cohorts of workers (young men and young to middle-aged women; see on this Krusell et al. 2010 and Elsby et al. 2013).

The chapter is organized as follows: section 2 presents the literature review on the topic, section 3 introduces the main issues that undermine the typical estimation methods; section 4 describes our model and the estimation process, section 5 presents our results and section 6 concludes.

3.2 Literature Review

This work encompasses two separate branches of literature: one discussing the estimation of potential unobserved variables, and one which tries to quantify the economic impact of the structural changes in the composition of the labor force.

The first strand of literature addresses the role that social phenomena such as aging, female labor participation, and schooling have for the structure of the labor market itself. Pissarides (1989) was among the first to highlight the importance

[4] This refers to the unreliability of traditional filters near the endpoints of the data set.

of population aging to economic research, in his innovative work analyzing the impact of demographic factors on the labor market. Among his theoretical findings, the robust result that population aging contributes to reduce unemployment is probably the most relevant with respect to our work. However, the topic had then been forgotten for more than a decade. It is in fact only in the early 2000s that Brsch-Supans (2001) work shed more light on the influence of demographic factors on unemployment. His models resulted in insights on how the intervention of the policy maker can mitigate the effects of aging on the labor market. The interest on the topic remained sporadic, and most of the works have focused on the general economic impact of population aging (see e.g. Bloom et al. 2011).

However, the interest in the role of aging in the labor market behavior has recently renewed. Hall (2014) showed that about half of the 2007-14 decline in labor-force participation is due to the aging of the population as the baby-boom generation retired. Similarly, Cline and Nolan (2014), using a simple regression model for participation rates found that demographic factors can account for up to two thirds of the changes in labor force participation in the US (mostly due to population aging). Surprisingly, many policy discussions do consider this issue, while only few academic studies underline the importance of this cause. In addition, existing measures to estimate trend employment and labor force participation rates do not have the ability to fully capture the impact of demographic changes.

On the other side, the discussion has been fervent in the development of innovative methods to estimate potential variables. In this literature, two very different approaches co-exist: we have fully-specified structural models, mainly developed by supranational institutions like the IMF or the European Commission (Havik et al., 2014), and which deliver precise estimates but that tend to be strongly model-dependent, and research papers that dig into the statistical methodology to improve the ways to deal with the non-observability of potential variables (for a full literature review on the different estimation method techniques, see Cotis et al. 2004).

Because of the relevance of the estimation of the NAIRU for the policy maker, the literature on this topic is vast; we here focus more narrowly on the use of state-space modeling to estimate potential variables. We have to highlight as the usage smoothing techniques to estimate the NAIRU has been wide in the last few years:

see Schumacher (2008), Fitzenberger et al. (2007), Apel and Jansson (1999ab), Laubach (2001), Fabiani and Mestre (2000, 2004) among others. In particular, the state-space framework is among the most commonly used: Basistha and Startz (2004) and Staiger et al. (1997) estimate the US NAIRU in different state-space frameworks, while Greenslade et al. (2003) focus on the UK in a slightly different statistical framework.

There have already been some attempts to refine the estimation method of the NAIRU, with Kalman Filter techniques performing better than HP in this context (see the discussion in Borio et al., 2012). At the same time, we are not the first ones to try to estimate long-term trend employment (Carone, 2005, develops a two-step methodology to make projections on the labor supply of the 25 EU member states), to underline the link between labor force participation and economic variables (Daly 2007; Monacelli et al. 2011; Petrosky-Nadeau 2014; Garn 2015; and Miao et al. 2016) or to disaggregate potential to retrieve more precise estimates (Fleischman and Roberts, 2011). The innovation behind our work is to bring together these branches of literature, and to unite them into a unique estimate for the long-term labor force, incorporating the structural brakes of the population via the estimation by cohort and gender and filtering out some proxies for the financial cycle.

3.3 Labor Market and Demography

Population aging is the result of a lower fertility rate - especially in developed countries, see Figure (3.1) - combined with a higher life expectancy - in particular that of men is rapidly catching up with that of women, Figure (3.2). Moreover, with "baby boomers" reaching the age of retirement, the ratio between workers and pensioners is falling down. This trend is putting increasing pressure on pension systems in most industrialized countries, and constraining the governments fiscal stance.

Another important factor affecting the dynamics of employment is the steadily increasing participation of women in the job market. This trend has at least two important consequences: one is the direct increase of labor supply, which contributes to enlarge the labor force and the level of employment. Not surpris-

Figure 3.1: Women fertility rate

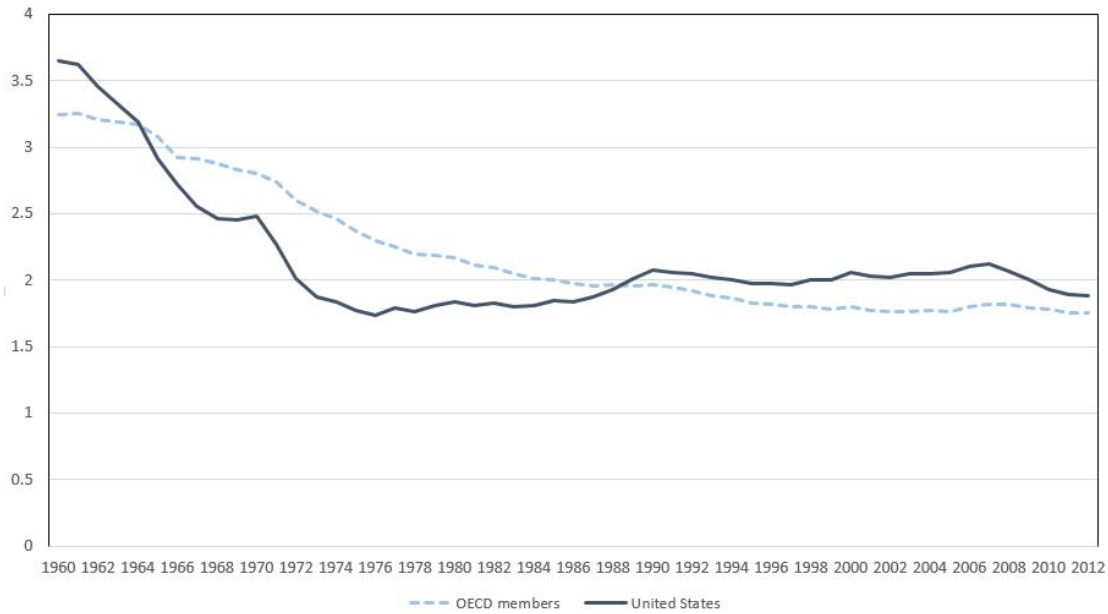
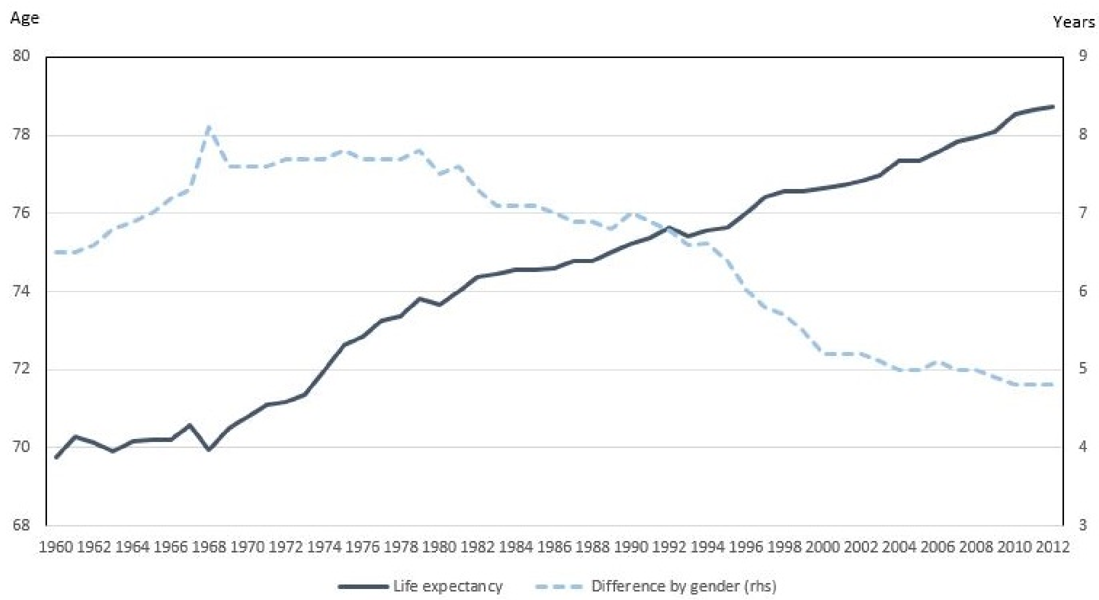
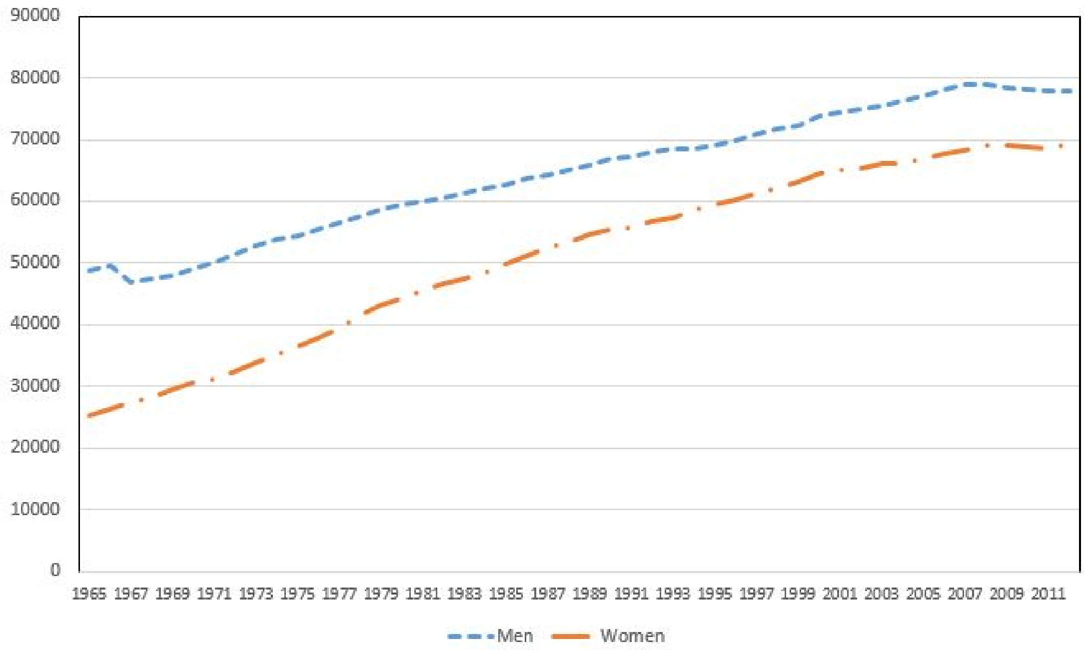


Figure 3.2: Life expectancy and gender difference, United States



ingly, the women employment kept on increasing during the recent global economic downturn (and decreased only slightly, right after the crisis, see Figure (3.3)).

Figure 3.3: Labor force, United States



The second and more recent phenomenon is strongly related to periods of crisis: when one member of the family unit remains unemployed, the other - often a woman - is likely to enter the labor force (Eltsby et al., 2013).

The increasing level of education is also contributing to modify the composition of employment, in particular for the youngest cohorts: secondary education and university delay the entrance of young workers into the labor market (CEA 2016). Our methodology, which estimates the potential level of employment for single cohorts, helps account for all these phenomena.

When analyzing the labor market, researchers tend to underestimate the role of demographic changes in modifying the composition of the labor force. The process of population aging creates a long term dynamic which - together with migrations - is deeply modifying the structure of the labor force.

In addition, the different population cohorts are affected by this process in a different way. It is reasonable that in younger cohorts potential employment reduced because of schooling and demography, while in the middle cohorts potential employment increased because of baby-boomers aging. The most common estimation methods fail to account for these long-term dynamics, and their estimates may be biased.

Instead, by estimating the potential employment within the single cohorts, we account for many of the causes of the long-term transition that are affecting the labor market: population aging, low fertility rates, migration, and schooling. For instance, the equilibrium measure of the male labor force in the 19-24 years old cohort will slowly embody the reduction of both population and people actively looking for a job due to schooling and fertility. Furthermore, the business cycle and other structural dynamics should differently affect the probability to find a job for agents in different cohorts (Krusell et al. 2010; Elsby et al. 2013). Once we aggregate all cohorts in our final estimation, the method will incorporate the long-term population dynamics in the results, with a large gain both in the precision of the identification of the potential employment, and in the ability of the model to produce reliable projections of the future employment.

3.4 Methodology

This study aims to increase the accuracy of potential employment estimates. In order to do so, we need to employ a methodology suitable to correctly disentangle the cyclical component of a variable from its trend. Three techniques are commonly used by scholars and policy makers to achieve this goal.[5]

A first strand of literature adopts a purely statistical approach to estimate a reduced form relation, as the Hodrick-Prescott filter or the estimation of time-series regressions. A second group of studies examine the equilibrium - or full-utilisation - level of the variables estimating calibrated theoretical models. In these models the extent of the gap between the equilibrium and the actual level of the variables strictly depends on the presence of frictions, or rationing, in the economic system.

[5]For an extensive discussion of the evolution of the methodologies to disentangle potential and cyclical components of economic variables - and on the concept of "potential" and "cyclical" itself - see, among others, Borio et al. (2012, 2014).

A third set of researches is a compromise between these two approaches: while they allow the potential component to be estimated with statistical techniques, they also rely on theoretical recommendations to impose certain identifying constraints on the path of the estimated variable.[6]

We have to highlight as the purely statistical approach and the model approach suffer from crucial problems. On the one hand, in theoretical model the frictions are considered exogenous, and hence cannot be a target measure for the policy-makers. As a consequence, different assumptions on the presence and nature of these frictions lead to different results (Borio et al. 2014).

On the other, the reduced form univariate model (such as the HP filter, the Baxter and King filter, or other unobserved component methods) suffers from the well-known end-point problem: the reliability of the end-of-the-sample estimates is limited. This concern affects the usefulness of the results for real-time analysis or policy decisions. Moreover, in the most used technique - the HP filter - the amplitude of the frequency is exogenously set by the researcher.

For these reasons Borio et al. (2012) proposed an alternative estimation process, using a refined version of the Kalman Filter (KF).[7] The higher accuracy of the KF with respect to the other procedures listed above is ensured by two different facts. First, the frequencies are no longer set exogenously by the researcher, but computed by the filter itself with an estimation update algorithm that enlarges the convergence speed to the true signal-to-noise ratio. The second is the possibility to make the filter better fit the data, by adding some additional explanatory variables - that we will call conditioning variables - to the estimation process.

This procedure assures transparency in the estimation methodology and simplicity compared with structural models. The explanatory variables that lack statistical significance are excluded from the estimated model since they fail in helping the filter to rule out the cyclical component. This can be tested with a standard *t-statistic*. Therefore, it is the model itself suggesting which variables are significant and have therefore to be included.

We include in the estimated model some proxies for the financial cycle. The

[6]This methodology has been widely used for the estimation of the NAIRU, see Turner et al. (2001).

[7]Unobservable Components Models have been widely used in economics since Harvey (1989) contribution

rationale behind this is that the presence of financial variables can significantly improve the identification, provided the statistical significance of the variables embedded in the model. This hypothesis is supported by recent studies on the role of financial cycle in determining labor market movements.

As Monacelli et al. (2011), Petrosky-Nadeau (2014), and Garn (2015) point out, changes in credit conditions affect the bargaining power of firms. According to these models, firms will borrow more when the credit conditions are favorable, increasing their bargaining power in the labor market. This, in turn, will increase their willingness to hire. Vice versa, when credit conditions worsen, firms' willingness to hire will decrease. This result is confirmed by studies like Miao et al. (2016), where the collapse of a credit bubble produces high and persistent unemployment.

In addition, as suggested by the Okun's law, any shocks affecting the business cycle have the ability to modify the level of (un)employment. Therefore, the documented relation between credit and output (Borio 2014) also spreads to the labor market.

Looking at our model in detail, the Kalman filter can be represented by a state-space model. This model will account for a *state equation* of the form:

$$\Delta e_t^* = \beta_1 \Delta e_{t-1}^* + \epsilon_{1,t} \tag{3.1}$$

and a *measurement equation*:

$$e_t = e_t^* + \beta_2(e_{t-1} - e_{t-1}^*) + \gamma' x_t + \epsilon_{2,t} \tag{3.2}$$

where $e_t = ln(E_t)$ represents the log of employment, e^* is its potential level, and Δe_t^* is its cyclical component. $\epsilon_{1,t}$ and $\epsilon_{2,t}$ are normally and independently distributed errors with zero mean and variance σ_1^2 and σ_2^2.[8] Finally, x_t is the vector including our the proxies for the financial and the business cycles - the interest

[8]For robustness, we also perform the estimation including a further lag in equation (3.1), which takes then the form:

$$\Delta e_t^* = \beta_1 \Delta e_{t-1}^* + \beta_3 \Delta e_{t-2}^* + \epsilon_{1,t}$$

To check whether β_1 is able to capture the whole persistence of the model. The results do not change substantially, an the beta is not significant.

rate (as proxy for the monetary policy), the inflation rate, the potential output, and the credit-to-gdp ratio.[9]

The KF jointly minimises the squared residuals in (3.1) and (3.2), since it calculates the least squares forecasts for the variables in the model. As a result, the solution for e^* will depend on the ratio of the two variances, which is called *signal-to-noise ratio* and is defined as:

$$\lambda_1 = \frac{\sigma_1^2}{\sigma_2^2} \tag{3.3}$$

The equations play a different role. Equation (3.1) defines the growth of the variable's potential level as an AR(1), with a persistence determined by β_1. Equation (3.3) allows for protracted one-sided deviation of potential employment estimates from the level of actual employment.

Equation (3.2) is the measurement equation. It anchors potential to actual employment, imposing that their difference is is described by an AR(1) process where a vector of proxy for the financial and the business cycles x_t carries significant information. The idea behind equation (3.2) is that the cyclical component, which fills the gap between the trend component of the variable and its actual value, is more than a normally distributed error term, as supposed by filters such as the HP (Borio et al. 2014). As a consequence, in this state space model the measurement error has a well-defined behavior that might be better identified thanks to the conditioning variables included in the vector x_t.

Equation (3.3) determines the relative variability of the estimated employment equilibrium level, setting the extent to which the potential employment is anchored to its actual level. When λ_1 becomes very large, our equilibrium level of employment will approximate a linear trend. If λ_1 gets close to zero, there will be no difference between the estimated trend and the actual measure. In our exercise, the value of λ_1 is settled equal to 100 for the HP filter (which we use as a benchmark), while we restrict the signal-to-noise in the Kalman Filter, $\lambda_1 = \frac{\sigma_1^2}{\sigma_2^2}$, to be:

[9] For further details on variables source and coverage, see the Appendix A. On the relevance of potential output for business see Borio et al. 2012.

$$\frac{var(e_t - e^*_{(hp),t})}{var(\Delta^2 e^*_{(hp),t})} = \frac{var(e_t - e^*_{(kf),t})}{var(\Delta^2 e^*_{(kf),t})} \tag{3.4}$$

where $e^*_{(hp),t}$ and $e^*_{(kf),t}$ are the potential output from the HP filtering and the Kalman Filter, respectively. In this way we preserve the frequencies of HP and Kalman estimations to be the same.

We estimate our model for each age-gender cohort of US employment. Our dataset uses OECD data on employment for the United States in the period 1960-2015 (extraction: August 2016). The data are annual and divided by gender and ages classes (the length of a class is 5 years, starting with the 19-24 and ending with the 65 and over).

In order to obtain the most parsimonious model, the choice of the variables to insert in equation (3.2) follows a *general-to-particular* procedure: we start by testing the significance of the variables (potential output, inflation rate, monetary policy, and government primary balance) in each cohort, which we considered meaningful to detect the cyclical component of the employment, and their three lags (results are available upon request). It is then the model itself to suggest the eventual statistical significance of the various variables included. The variables included in the final specification are reported in table (3.1) - that will be discussed in the following sections -, with the respective estimated coefficients and $t - statistic$.

We want to highlight two crucial aspects of our methodology. First, - as expected - different cohorts present a different behavior of the trend component. For instance, the largest part of male cohorts (25-29 to 50-54) rose steadily until the 90s and then seem to stabilize. Moreover, also the cyclical frequency varies across cohorts. Therefore, the signal-to-noise ratio has to continuously adjust.

Second, building on Borio et al. (2012) the model is estimated with Bayesian techniques. We employed gamma distributions as priors for all coefficients, and inverted gamma distribution for the error terms. In our priors autoregressive coefficients are restricted to lie between 0 and 0.99, while coefficients for the conditioning variables is restricted to be positive. In order to assure a sufficient persistence for both the trend and cyclical components the prior means for the autoregressive coefficients is fixed, respectively, to 0.85 for the trend and 0.6 for the cyclical (both with a prior variance of 0.6), while for conditioning variables we opted for a prior mean of 0.3 (with a variance of 0.3).

We then aggregate the estimated potential and cyclical components of employment. The aggregation is done with a weighted average, where the weights are the shares of single cohorts in total employment - i.e., cohort employment over whole employment. The final result of these estimates is a new series for employment, with the cyclical component filtered out and able to account for the demographic factors modifying the composition of the labor force.

It is important to highlight that the conditioning variables should have a stable mean. This characteristic is rarely present in economic time series, which also tend to show a high degree of cyclicality. Following Borio et al. (2012), we therefore decided to demean the conditioning variables via Cesàro means,[10] which increase the rate of convergence of our model.

3.5 Results

3.5.1 Cohorts Estimations

In this section we report the main figures of our results. The Appendix contains all the estimation for the final model for each cohort, i.e., the complete list of results for the second step of our methodology, plus some additional robustness experiments.

Table (3.1) reports the statistical significance of our proxies. After the *general-to-specific* procedure we included, in different cohorts, the credit over gdp cr_t and the lag of inflation inf_{t-1}.[11] In addition, table (3.1) contains also the results for the autoregressive coefficients β_1 and β_2 (respectively for trend and cyclical components). For each coefficient, we provide both the estimated coefficients and their respective t-statistics in each cohort. As Borio et al. (2014) underline, the

[10]Named after Ernesto Cesàro, who proved that if a sequence of numbers converge to a constant - the mean - the sequence of arithmetic means taken over the first n first elements also converge to the same constant.

[11]In the variable-by-variable estimations also the lagged value of credit was significant. However, when we estimated the final model it always result not significant, and therefore we excluded it in the final specification

significance of a variable implies not only that this variable is correlated with the employment, but also with the *frequencies* implicitly set by the scaling factor.

Some considerations are in a row. First, male and female employment seems to be very reactive to credit conditions. Among the cohorts analyzed, only the 50-54 for male does not show significant coefficients (the critical values are 1.298 at 10%, 1.675 at 5%, and 2.400 at 1%). An increase in the credit over gdp level, which is a proxy of the presence of a boom in the credit market, is correlated with a higher (cyclical) employment level. Male and female prime-aged cohorts (15-19 and 20-24) are the ones with highest coefficients, followed by the 55-59 cohort for male and 65+ for female: not surprisingly, among males, the more affected by contemporaneous credit conditions are the prime aged - the ones that have to decide whether to enter the labor market or not, and those that have to decide whether to retire or not (cohort 55-59). These results are in line with the aforementioned literature on the role of credit in the behavior of cyclical employment (Monacelli et al. 2011; Petrosky-Nadeau 2014 and Garn 2015; Miao 2016).

Business cycle conditions, approximated by the lagged inflation rate, in general do not play a significant role in determining cyclical employment. We observe significant coefficients in the cohorts 55-59 and 40-44 for males; and 50-54, 30-34, and 15-19 for females: this indicates that middle-aged employee are more affected by the movements of the business cycle.

The estimations for β_1 lie between 0.87 and 0.97 and are strongly significant: as expected, the trend component is very persistent. β_2 coefficient is more variable and not always significant. However as highlighted by Borio et al. (2014) the presence of an autoregressive term enhance the estimation robustness, while does not modify the punctual result.

Figure (3.4) shows the results we obtain for male employment in the age cohort 15-19[12]. Panel (a) shows the behavior of actual employment (blue line) compared with the KF estimated potential (red, dotted line) and the HP estimated potential

[12]This has to be considered as an example of the gains of our methodology with respect to other techniques, such as the HP filter. As we said above, our full estimates are reported in the Appendix. We decided to not report the full set of estimates here both because the they present a similar behavior among cohorts, and because our final goal is to discuss the most relevant characteristics of our methodology. The final estimated series of employment is however obviously available upon request.

level (yellow, pointed line). Panel (b) compares the HP cyclical component (red, dotted line) with the KF cyclical component (blue line). Panel (c) and (d) report the KF - Panel (c) - and HP - Panel (d) - cyclical components with a blue line, and the relative confidence bands, with a red pointed line.

The KF filter estimation of the cyclical component shows much narrower confidence intervals if compared to the HP estimates. This is a signal of an "identification gain", which is the result of the larger information set that enters the KF thanks to the proxies for the financial and business cycle.[13]

We can observe that the KF estimates a larger cyclical loss with respect to the HP during the last economic crises. Such result is in favor of a lower role of the 2008 recession in diminishing *potential* employment, while supports a larger impact on cyclical employment. In particular, the larger fall in cyclical employment is related to the inclusion of credit-over-GDP in the specification: excluding this from the specification the gap between HP and KF estimations reduces. This finding is similar to the one of Borio et al. (2012), where the inclusion of credit-to-gdp ratio among in the model increase the output gap in the Great Recession. Therefore, without considering the behavior of financial markets models fail in identify cyclical and trend components, overestimating the role of the recessions in influencing potential variables.

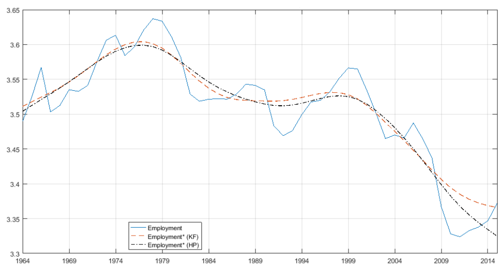
3.5.2 Aggregate Estimations

In this section we present the estimations of potential and cyclical employment obtained by aggregating the KF estimation results in each of the single cohorts. From now on, we will refer to these measures as "aggregate KF".

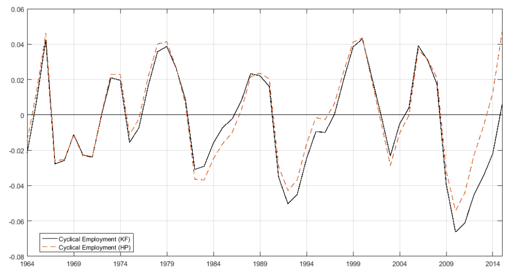
We divide this section into two parts: in the first, we discuss the performance of the aggregate KF compared with: a KF on total employment, to the potential level of employment as estimated by the OECD, and to "aggregate cohort" HP. In the second, we present the gains in terms of real-time estimation, i.e. to what extent

[13]In the HP filter the cyclical component is imposed to be an erratic term. Specifically, in our model the estimated standard error of the cyclical component is 0.0271 for the HP and 0.0186 for the KF. This imply a lower indeterminacy of the cyclical component. In addition, the R^2 of the HP is 0.884 while is 0.945 for the KF, supporting that the latter explains a larger share of the employment variation.

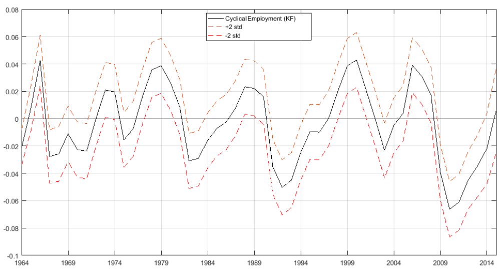
Figure 3.4: Male Employment cohort age 15-19 - US



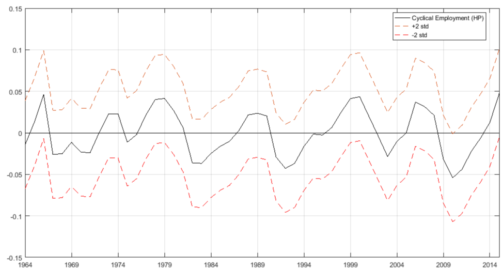
(a) Actual and Potential Employment



(b) KF and HP Cyclical Employment



(c) KF Cyclical Emp and Confidence Bands



(d) HP Cyclical Emp and Confidence Bands

Table 3.1: Regression results: significance of explanatory variables

Male Employment											
Variable	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65+
β_1	0.871***	0.889***	0.903***	0.954***	0.945***	0.928***	0.946***	0.954***	0.911***	0.950***	0.974***
	(14.348)	(14.210)	(2.455)	(2.863)	(3.574)	(2.919)	(3.732)	(4.100)	(6.768)	(2.856)	(4.225)
β_2	0.237*	0.315**	0.227	0.652	0.465	0.604	0.242	0.394**	0.286*	0.614	0.335
	(1.478)	(1.786)	(1.100)	(0.163)	(1.275)	(0.146)	(0.151)	(2.000)	(1.305)	(0.154)	(1.720)
cr_t	1.283***	1.014***	0.311*	0.807***	0.652***	0.898***	0.462***	0.114	0.882***	0.490**	0.593**
	(3.047)	(2.799)	(1.436)	(3.735)	(3.590)	(4.372)	(2.492)	(0.715)	(6.312)	(2.217)	(2.230)
inf_{t-1}	0.222	0.196	0.144	0.236	0.003	0.532***	0.108	0.163	0.608***	0.265	0.048
	(0.471)	(0.505)	(0.638)	(1.056)	(0.017)	(2.509)	(0.560)	(1.030)	(4.140)	(1.125)	(0.169)
Female Employment											
Variable	15-19	20-24	25-29	30-34	35-39	40-44	45-49	50-54	55-59	60-64	65+
β_1	0.905***	0.950***	0.946***	0.901***	0.948***	0.900***	0.954***	0.933***	0.954***	0.948***	0.949***
	(17.326)	(5.534)	(2.839)	(5.890)	(4.652)	(3.952)	(3.966)	(3.378)	(4.012)	(3.076)	(8.508)
β_2	0.237*	0.685***	0.586***	0.260	0.262	0.614	0.700	0.373**	0.429*	0.797***	0.682***
	(1.586)	(4.061)	(3.417)	(1.092)	(1.283)	(0.153)	(0.175)	(2.148)	(1.377)	(3.882)	(3.624)
cr_t	1.527***	0.910***	0.623***	0.946***	0.982***	0.486***	0.331**	0.621***	0.383**	0.402**	0.991***
	(4.503)	(4.948)	(3.416)	(4.860)	(5.353)	(2.514)	(2.056)	(3.543)	(2.130)	(2.371)	(3.736)
inf_{t-1}	0.556*	0.229	0.173	0.769	0.196***	0.045	0.095	0.714***	0.065	0.0638	0.1585
	(1.478)	(1.220)	(0.875)	(4.149)	(0.994)	(0.218)	(0.551)	(4.019)	(0.347)	(0.361)	(0.617)

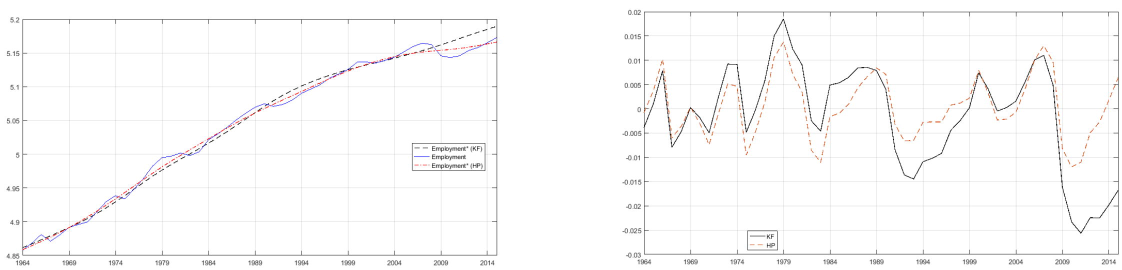
*Note: *10%, **5%, ***1% significance level. For each cohort is repoted the explanatory variables estimated coefficient and the corrispective t-statistic in parenthesis.

the methodology helps in dealing with the end-point problem, and in explaining the behavior of other macro-variables. All the figures report logs of estimated and actual values.

Figures (3.5) and (3.6) show the results for the aggregate KF compared with the aggregate HP and the OECD potential (and cylical) employment measures, respectively.[14] In both figures, the left panel reports the potential estimation (red pointed line the aggregate HP and OECD, black dotted line the KF), with the actual levels (blue, continuous line). The right panel reports the cyclical components without the confidence bands: the continuous black line is the estimation of the aggregate KF, while the red pointed line is the aggregate HP (Figure (3.5)) and OECD measure (Figure 3.6).

The the two estimations have patterns similar to the one in the cohort 15-19 analyzed above. Compared to the aggregate HP, the aggregate KF estimates a larger loss in the cyclical employment during the last economic crisis and while the there are larger positive changes until 1991.

Figure 3.5: Aggregate KF VS Aggregate HP

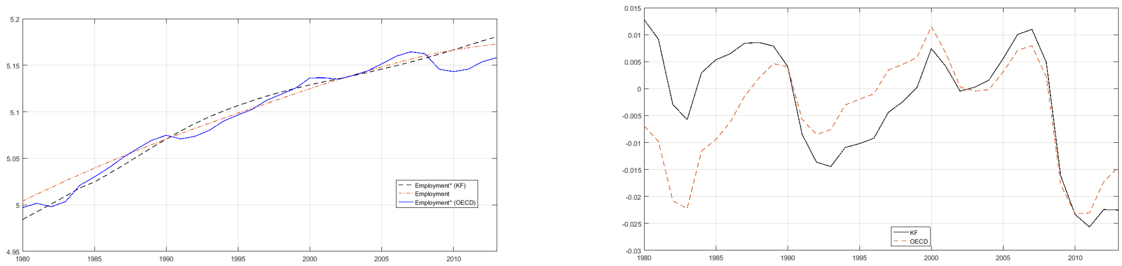


(a) Actual and Potential Employment (KF and HP) (b) Aggregate KF and HP Cyclical Employment

Our methodology also performs well with respect to the OECD potential employment estimations (Figure (3.6)). In particular, the gap among the two measures during the Great Recession is tiny compared to the one with the aggregate HP. On the contrary, until 1990, the OECD measure estimates a large cyclical

[14]OECD potential (and cyclical) employment components are based on a mixed model, in which a NAIRU estimated with a statistical approach is then used in a model to estimate potential output and the other potential variables - see the Statistical Annex to the OECD Economic Outlook. This methodology has been recently modified, and data for the US covers the period 1980-2013 - more recently data only back to 1985.

Figure 3.6: Aggregate KF VS OECD Potential Employment

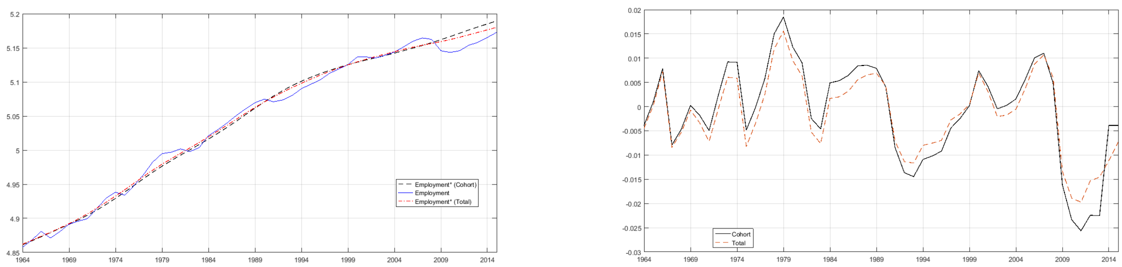


(a) Actual and Potential Emp (KF and OECD) (b) Aggregate KF and OECD Cyclical Employment

employment loss, larger than the ones computed with both the aggregate KF and the aggregate HP.[15]

Therefore, our estimations lie somewhere in between the ones provided by the other methodologies investigated. They perform reasonably well both at the end-point - which is a concern for the HP - with estimates similar to the OECD procedure, and the start-point, with estimates similar to the HP.

Figure 3.7: Aggregate KF VS Simple KF



(a) Actual and Potential Employment (b) Aggregate and Simple KF Cyclical Employment

The aggregate KF estimates (Figure (3.7)) show a smoother path with respect to the standard KF ones applied to the aggregate employment. In particular,

[15]This difference is wide: the OECD measure estimates a negative cyclical employment between 1980 and 1987 (aggregate HP: 1982-1985; aggregate KF 1982-1984), with a trough of -5% (aggregate HP: -2.65%; aggregate KF: -1.1%. This, together with the fact that the procedure has been recently revised, casts doubts on the validity of these estimates - especially in the early part of the sample.

the cyclical employment (Figure (3.7), panel (b)) for the aggregate KF has a smoother path with lower oscillations. As a matter of fact, the cyclical effect of the financial crises is larger in the non-aggregate KF estimation. This result implies that without taking into account the population dynamics the filter is more sensible to deep recessions.

This intuition is confirmed by analyzing to what extent cyclical employment rate explains macro-variables behavior. This exercise consists in running some auxiliary regressions where growth rates of output (nominal and real) and inflation are regressed on the cyclical component of employment estimated with different techniques: the aggregate HP, the Kalman filter on aggregate data, the aggregate Kalman filter, and the OECD estimations (see table (3.2)). We employ two different estimation frameworks, a simple OLS and an AR(2), and we report estimated coefficients, $p-values$, and R^2.

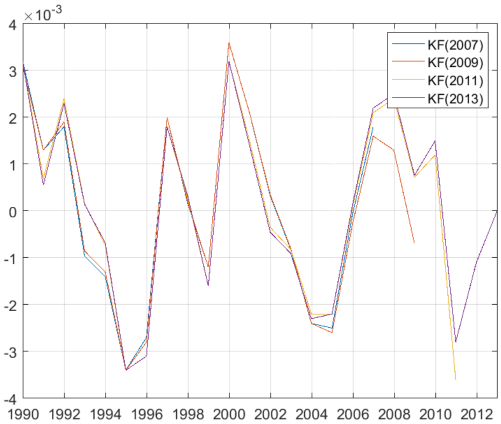
The regressions containing the series estimated via Kalman filter have larger and significant coefficients for nominal GDP and inflation, while they are smaller - but significant - for real GDP. The R^2 suggest that the regression employing the aggregate KF explains a larger portion of the variation of the dependent variable compared with the other methodologies, the only exception being the regression for output with AR(2) components, where the simple KF has a larger R^2: however, in this case the coefficient for the simple KF is not significant, while the one for the aggregate KF is significant at 10%.

Another important feature of our estimation method is that it seems to be less subject to *ex-post* revisions, and to the addition of new data with respect to simple filtering. Figure (3.8) reports the in-sample estimations of the cyclical component of employment with various end-points. We present the results for three methodologies: the aggregate KF, the aggregate HP, and the OECD model. We evaluate the series at four different end-points: 2007, 2009, 2011, 2013 (we excluded the last two years because of missing observations in OECD time series).[16]

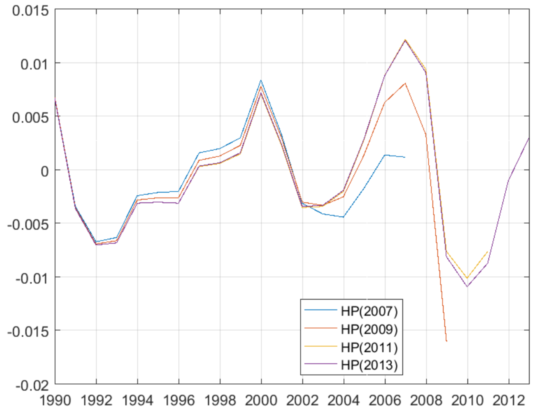
This exercise helps in evaluating the robustness of our model to the end-point problem and, consequently, in assessing its ability to provide reasonable real-time

[16]Unfortunately, as we highlighted above, the OECD model has been subject to deep modifications both in 2008 and in 2012: we were not able to retrieve the contemporaneous estimations for cyclical employment in 2007.

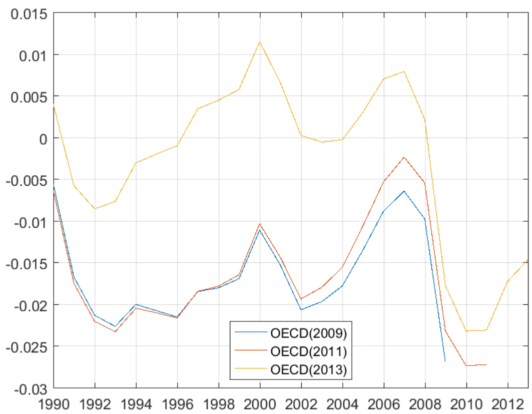
Figure 3.8: Cyclical Employment Estimation at Different Horizons



(a) Aggregate KF



(b) Aggregate HP



(c) OECD Model

Table 3.2: Regression results: significance of cyclical components

	Simple Regression				AR(2) Regression			
Variable	Aggregate KF	Simple KF	Aggregate HP	OECD	Aggregate KF	Simple KF	Aggregate HP	OECD
gdp	0.467*	0.578	0.616	-0.365	0.594*	0.817	0.950	-0.349
$p-value$	(0.096)	(0.116)	(0.185)	(0.363)	(0.077)	(0.071)	(0.105)	(0.317)
R^2	[0.054]	[0.048]	[0.035]	[0.025]	[0.177]	[0.182]	[0.170]	[0.076]
$ngdp$	1.614***	1.836***	1.226*	0.490	1.447***	1.487***	1.333**	0.407
$p-value$	(0.000)	(0.000)	(0.055)	(0.274)	(0.004)	(0.009)	(0.016)	(0.434)
R^2	[0.334]	[0.253]	[0.071]	[0.037]	[0.560]	[0.551]	[0.534]	[0.153]
inf	1.454***	1.723***	1.189	1.001*	1.409***	1.403**	1.242***	0.706
$p-value$	(0.000)	(0.000)	(0.051)	(0.058)	(0.000)	(0.016)	(0.006)	(0.232)
R^2	[0.297]	[0.229]	[0.073]	[0.107]	[0.747]	[0.739]	[0.723]	[0.211]

* Note: *10%, **5%, ***1% significance level. The table reports estimated coefficients, standard errors (parenthesis) and R^2 (squared brackets).

estimations. As Figure (3.8) illustrates, our model seems to perform well. Especially around the end of the sample, the aggregate KF is subject to lower *ex-post* revisions than the aggregate HP and the OECD model. In particular regarding the latter methodology, Figure (3.8) highlights the major issue of a model-based approach to potential/cycle components identification: when the model is revised, the estimates vary significantly.

3.6 Conclusion

In this chapter we derive an innovative method for filtering the cyclical component out of employment. The estimation procedure consists of a state-space model in which the measurement equation is augmented to include some proxies for the cycle. We performed this analysis by age cohort and gender, in order to be able to account for the demographic factors that are currently modifying the composition of the labor force (population aging, migration, increasing female participation). This allowed us to obtain an estimated series that is free of influence from both business cycle and demographic factors.

We compared the results of our methodology with the most important alternative: a simple Kalman Filter on aggregate employment, the HP filter, and the OECD cyclical employment. Our model outperforms this "traditional" methods under several dimensions. First, as the analysis of the R^2 obtained regressing these variables on the cyclical employment suggests, our aggregate Kalman Filter explains a larger portion of the variation of macro-variables growth compared with other methodologies. Second, it is more robust to in-time and ex-post revisions. Third, the indeterminacy of the cyclical component reduces, since the standard errors are generally smaller compared with the HP ones.

In addition, the proxy for the financial business cycle, the credit-over-gdp ratio, is significant in many of the cohorts analyzed. This supports a recent theoretical literature (Monacelli et al. 2011; Petrosky-Nadeau 2014; and Garn 2015) which suggest that the financial cycle can directly influence the labor market, and the job creation in particular.

The middle-aged cohort responds to variation in inflation - which we interpret as a proxy for the business cycle. This result is in line with most studies on the

Phillips curve and the NAIRU, in finding that inflation contributes to explain the cyclical behavior of employment. Finally, our results suggest that the levels of employment among the youngest cohorts are the most closely related with the financial and business cycle. This last result is of particular interest for studies on labor market behavior, and it should be better investigated by future researches.

The results have to be refined in two directions. The first is methodological: in this framework, projecting the variables trend for a long period is complex and time-consuming. However, one of the aims of estimating unobserved components is to forecast future behaviors of variables. This claims for a methodological improvement obtain forecasts of the components in directly through the model.

The second is theoretical. The fact that prime-aged are more sensible to credit conditions seems reasonable, but could be further investigated in a future research as the goal of an *ad hoc* investigation.

Bibliography

[1] Aaronson D., Mazumder B. and Schechter S., 2010, *What is Behind the Rise in Long-term Unemployment?*, Federal Reserve Bank of Chicago, 2Q/2010 Economic Perspectives.

[2] Apel, M. and Jansson, P., 1999a, *System Estimates of Potential Output and the NAIRU*, Empirical Economics, 24, 373-388.

[3] Apel, M. and Jansson, P., 1999b, *A Theory Consistent System Approach for Estimating Potential Output and the NAIRU*, Economics Letters, 64, 271-275.

[4] Ball L., 1997, *Disinflation and the NAIRU*, in Romer C. D. and Romer D. H. (eds.), *Reducing Inflation: Motivation and Strategy*, National Bureau of Economic Research, University of Chicago Press, 167-194.

[5] Basistha A. and Startz R., 2004, *Measuring the NAIRU with Reduced Uncertainty: A Multiple Indicator-Common Component APproach*, University of Washington, mimeo.

[6] Blanchard O., and Summers L. H., 1986, *Hysteresis and the European unemployment problem*, NBER Macroeconomics Annual 1986, MIT Press, 1, 15-90.

[7] Bloom D.E., Canning D. and Fink G., 2011, *Implications of Population Aging for Economic Growth*, PGDA Working Paper Series n. 64.

[8] Borio C., Disyatat P. and Juselius M., 2012, *Rethinking Potential Output: Embedding Information about the Financial Cycle*, BIS Working Papers, 404.

[9] Borio C., Disyatat P. and Juselius M., 2014, *A Parsimonious Approach to Incorporating Economic Information in Measures of Potential Output*, BIS Working Papers, 442.

[10] Brsch-Supan A., 2001, *Labor Market Effects of Population Aging*, NBER Working Paper n. 8640.

[11] Brckner M. and Pappa E., 2012, *Fiscal Expansion, Unemployment, and Labor Force Participation: Theory and Evidence*, International Economic Review 53, 4, 1205-1228.

[12] Burniaux J.M., Duval R. and Jaumotte F., 2004, *Coping with Ageing: A Dynamic Approach to Quantify the Impact of Alternative Policy Options on Future Labor Supply in OECD Countries*, OECD Economics Department Working Papers, n. 371.

[13] Butz P. and Ward M. P., 1979, *The Emergence of Countercyclical U.S. Fertility*, American Economic Review 69, 3, 318-328.

[14] Carone G., 2005, *Long-term Labour Force Projections for the 25 EU Member States: A Set of Data for Assessing the Economic Impact of Ageing*, European Commission Economic Papers, n. 235.

[15] Chen J., Kannan P., Loungani P. and Trehan B., 2011, *New Evidence on Cyclical and Structural Sources of Unemployment*, Federal Reserve Bank of San Francisco Working Paper Series, 2011-17.

[16] Cline W.R. and Nolan J., 2014, *Demographic Versus Cyclical Influences on US Labor Force Participation*, Peterson Institute for International Economics, Working Paper 2014-4.

[17] Council of Economic Advisers, 2016, *The Long Term Decline in Prime-age Male Labor Force Participation.*

[18] Cotis J., Elmeskov J. and Morougane A., 2004, *Estimates of potential output : benefits and pitfalls from a policy perspectives*, London : Centre for Economic Policy Research 2004, 35-60.

[19] Daly M., 2007, *Labor Force Participation and the Prospects for US Growth*, Federal Reserve Bank of San Francisco Economic Letter, 2007-33.

[20] Dickens W., 2009, *A New Method for Estimating Time Variation in the NAIRU*, in Fuher J., Kodrzycki Y. K., Sneddon Little J. and Olivei G. P. (eds.), *Undestanding Inflation and the Implications for Monetary Policy. A Phillips Curve Retrospective*, MIT Press, 205-228.

[21] Elsby M. W. L., Hobijn B., Sahin A. and Valletta R. G., 2011, *The Labor Market in the Great Recession: an Update*, Federal Reserve Bank of San Francisco Working Paper Series, 2011-29.

[22] Elsby M., Hobijm B. and ahin A., 2013, *On the Importance of the Participation Margin for Market Fluctuations*, Federal Reserve Bank of San Francisco Working Paper Series, 2013-05.

[23] Fabiani, S. and Mestre, R., 2000, *Alternative Measures of the NAIRU in the Euro Area: Estimates and Assessment*, ECB Working Paper 17.

[24] Fabiani, S. and Mestre, R., 2004, *A System Approach for Measuring the Euro Area NAIRU*, Empirical Economics, 29, 311-341.

[25] Fitzenberger B., Franz W. and Bode O., 2007, *The Phillips Curve and NAIRU Revisited: New Estimates for Germany*, ZEW Centre for European Economic Research Discussion Paper 07-070.

[26] Fleischman C. and Roberts J., 2011, *From Many Series, One Cycle: Improved Estimates of the Business Cycle from a Multivariate Unobserved Components Model*, Federal Reserve Board Finance and Economics Discussion Series 2011-46.

[27] Garn J., 2015, *Borrowing Constraints, Collateral Fluctuations, and the Labor Market*, Journal of Economic Dynamics and Control, 57, 112-130.

[28] Goldstein J., Kreyenfeld M. Jasilioniene A. and rsal D. K., 2013, *Fertility Reactions to the "Great Recession" in Europe: Recent Evidence from Order-specific Data*, Demographic Research 29, 4, 85-104.

[29] Greenslade J. V., Pierse R. G. and Saleheen J., 2003, *A Kalman Filter Approach to Estimating the UK NAIRU*, Bank of England Working Paper n.179.

[30] Hall R.E., 2014, *High Discounts and High Unemployment*,NBER Working Paper.

[31] Harvey A. C., 1989, *Forecasting Structural Time Series Models and the Kalman Filter*, Cambridge University Press.

[32] Havik K., Mc Morrow K., Orlandi F., Planas C., Raciborski R., Roger W., Rossi A., Thum-Thysen A. and Vandermeulen V., 2014, *The Production Function Methodology for Calculating Potential Growth Rates and Output Gaps*, European Commission Economic Papers 535.

[33] Krusell P., Mukoyama T., Rogerson R. and ahin A., *Aggregate Labor Market Outcomes: The Roles of Choice and Chancel*, Quantitative Economics 1, 1, 97-127.

[34] Laubach, T., 2001, *Measuring the NAIRU: Evidence From Seven Economies*, Review of Economic and Statistics, 83, 218-231.

[35] Logeay C. and Tober S., 2006, *Hysteresis and the NAIRU in the Euro Area*, Scottish Journal of Political Economy, 53, 4, 409-429.

[36] Martin V., Hurn S., and Harris D., 2012, *Econometric Modelling with Time Series, Specification, Estimation and Testing*, Cambridge University Press.

[37] Massey D. S:, 2012, *Immigration and the Great Recession*, The Russel Sage Foundation.

[38] Miao J., Wang P. and Xu L., 2016, *Stock Market Bubbles and Unemployment*, Economic Theory, 61, 2, 273-307.

[39] Monacelli, T., Perotti R., and Trigari A., 2010, *Unemployment Fiscal Multipliers*, Journal of Monetary Economics 57, 5, 531-553.

[40] Monacelli, T., Quadrini V., and Trigari A., 2010, *Financial Markets and Unemployment*, NBER Working Paper N. 17389.

[41] rsal D. D. K. and Goldstein J. R., 2010, *The Increasing Importance of Economic Conditions on Fertility*, MPIDR Working Paper WP 2010-014.

[42] Passel J., Cohn D., Gonzalez-Barrera A., 2012, *Net Migration from Mexico Falls to Zero - and Perhaps Less*, Pew Research Center.

[43] Petrosky-Nadeau N., 2014, *Credit, Vacancies and Unemployment Fluctuations*,Review of Economic Dynamics, 17, 2, 191-205.

[44] Pissarides C.A., 1989, *Unemployment Consequences of an Aging Population*, European Economic Review, 33, 355-366.

[45] Scalone S., 2015, *Embedding Liquidity Information in Estimating Potential Output*, University of Verona, Department of Economics Working Papers, 2014-20.

[46] Sobotka T., Skirbekk V. and Philipov D., 2010, *Economic Recession and Fertility in the Developed World: A Literature Review*, Report for the European Commission, Vienna Institute of Demography.

[47] Schumacher C., 2008, *Measuring Uncertainty of the Euro Area NAIRU: Monte Carlo and Empirical Evidence for Alternative Confidence Intervals in a State- Space Framework*, Empirical Economics, 34, 357-379.

[48] Staiger D., Stock J. H. and Watson M. W., 1997, *How Precise are Estimates of the Natural Rate of Unemployment?*, in Romer C. D. and Romer D. H. (eds.), *Reducing Inflation: Motivation and Strategy*, National Bureau of Economic Research, University of Chicago Press, 195-246.

[49] Tafuro A., 2015, *The Effects of Fiscal Policy on Employment: An Analysis of Aggregate Evidence*, University Ca' Foscari of Venice, Dept. of Economics Research Paper Series 03/WP/2015.

[50] Turner D., Boone L., Giorno C., Meacci M., Rae D. and Richardson P., 2011, *Estimating the Structural Rate of Unemployment for the OECD Countries*, OECD Economic Studies, 33, 2011/II.

[51] Turrini A., 2013, *Fiscal Consolidation and Unemployment: does EPL Matter?: A Look at EU Countries*, IZA Journal of Labor Policy 2, 1, 1-18.

Chapter 4

Conclusions

The goal of this thesis was to investigate some important elements of financial stability. The analysis is conducted from three different perspectives and points of view, that correspond to the three chapters of this work, which partly retrace the study, research and work path of the author.

Contributions and innovations to the literature on financial stability come from each of the three chapters; however, despite bringing the analysis further, some issues in each topic remain open.

Chapter one is devoted to the analysis of the stability of the banking sector. It presents a machine learning technique to timely identify cases of financial distress, developed on a sample of relatively small European banks. The new and broader definition of distress allows to significantly enlarge the sample of positive events on which to estimate the model, while the new modelling technique ensures robust and reliable results. However, the work could be refined in at least two ways: first, the prediction horizon is relatively short (1-3 months), which undoubtedly limits the practical usefulness of the model; a longer prediction horizon would allor for a wider time frame for the supervisor to intervene, and potentially intercept or mitigate the cases of financial distress. Second, the composition of the dataset could be improved, in particular with respect to country-level macro variables. Givren the strong polarisation of the sample towards some jurisdictions (namely Germany, Austria and Italy), country variables like deficit over GDP ratio tend to become country proxies rather than indicators of the financial situation of a jurisdiction. In order to partly mitigate this issue, further developments of this work could include the inclusion of variables at the regional level (such as regional GDP, regional unemployment).

Chapter two presents a state-space methodology based on Borio et al. (2013) to build a "finance neutral" measure of potential output. We do so by introducing some proxies for the cycle in a state space framework, by augmenting the measurement equation. The results obtained with this methodology are more precise and robust with respect to the methods we use as benchmark, and we therefore confirm most of the results of Borio et al. (2013). Further research on this topic cannot overlook the creation of a fully specified model to better define the estimation framework.

Finally, chapter three proposes a two-step approach to estimate trend employment. While the methodological framework is not dissimilar from that of Chapter two (a state-space model with augmented measurement equation, to include proxies for the cycle), the scope of the investigation is different. The goal is in fact to retrieve a time series for long term employment, that would incorporate demographic dynamics that are often disregarded by traditional estimation methods: we conduct the analysis by gender and age-cohort, filtering out the cyclical component separately. Even in this third chapter there is room for a further expansion of the research: one might investigate why prime-aged workers are the most affected by credit conditions, by conducting an *ad hoc* empirical analysis, or incorporate a forward looking perspective directly in the estimation process.

Chapter 5

Appendices

5.1 Appendix A: List of the Variables

List of the variables used, including only the variables included in the final specification of the model:

- *Inflation rate*: consumer price index, United States, annual data 1960-2015, St. Louis Fed,

- *Interest rate*: monetary policy interest rate, 3-month treasury bill, United States, annual data 1960-2015, St. Louis Fed,

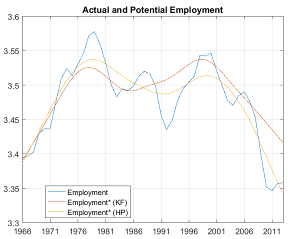
- *Labor force and population*: disaggregated for gender and age cohorts, United States, annual data 1960-2015, OECD-LFS statistics,

- *Credit to non-financial corporations*: credit to private non-financial sector (PNFS), United States, annual data 1960-2015, Bank for International Settlements.

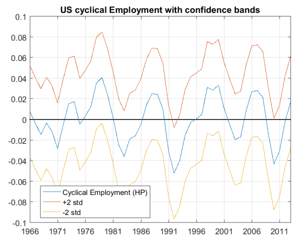
- *Output gap*: calculated as percent deviation of actual gdp from its potential, United States, annual data 1960-2015, St. Louis Fed.

5.2 Appendix B: Supplementary Results

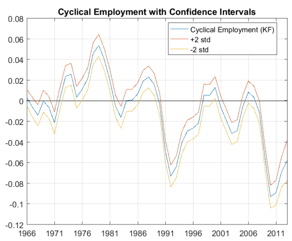
Figure 5.1: Results for females by age cohort 15-19 and 20-24



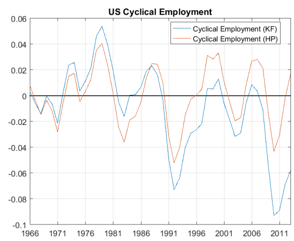
(a) 15-19 Actual and potential employment



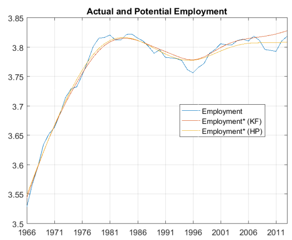
(b) 15-19 HP confidence bands



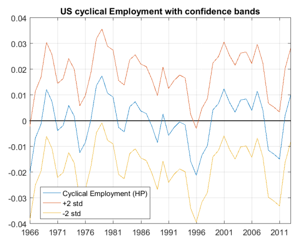
(c) 15-19 Actual and potential employment



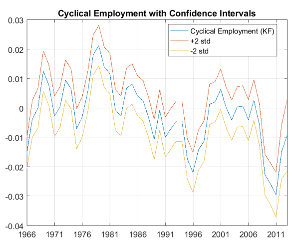
(d) 15-19 HP confidence bands



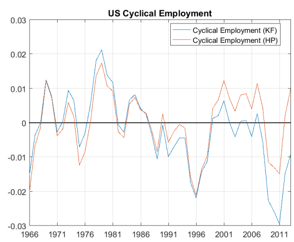
(e) 20-24 Actual and potential employment



(f) 20-24 HP confidence bands

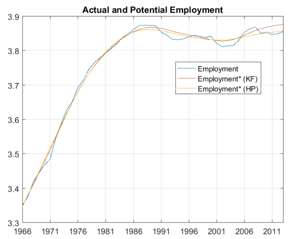


(g) 20-24 Actual and potential employment

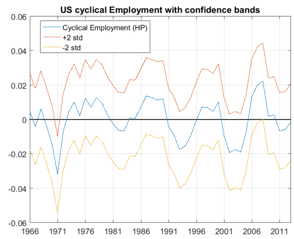


(h) 20-24 HP confidence bands

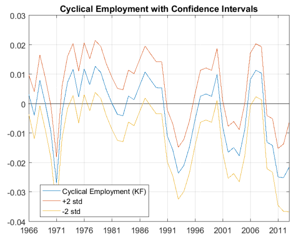
Figure 5.2: Results for females by age cohort 25-29 and 30-34



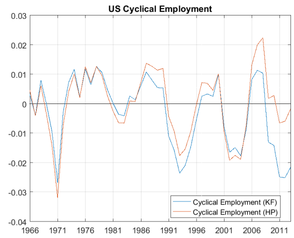
(a) 25-29 Actual and potential employment



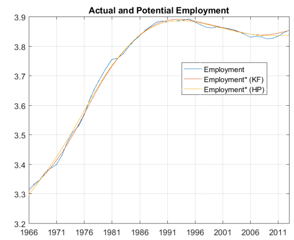
(b) 25-29 HP confidence bands



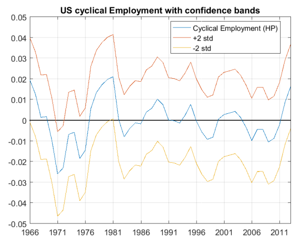
(c) 25-29 Actual and potential employment



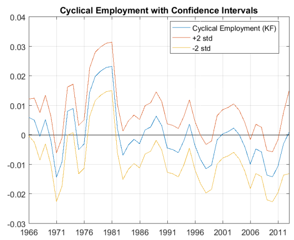
(d) 25-29 HP confidence bands



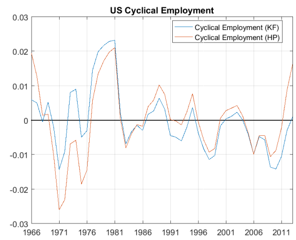
(e) 30-34 Actual and potential employment



(f) 30-34 HP confidence bands

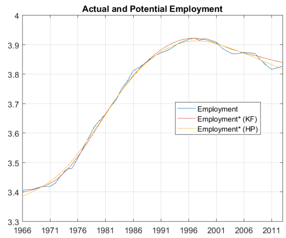


(g) 30-34 Actual and potential employment

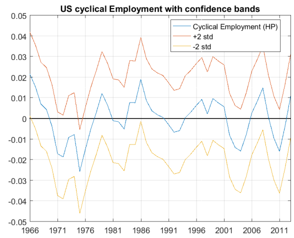


(h) 30-34 HP confidence bands

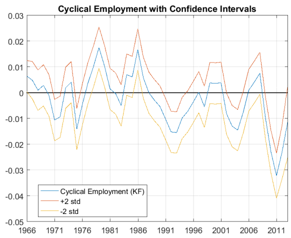
Figure 5.3: Results for females by age cohort 35-39 and 40-44



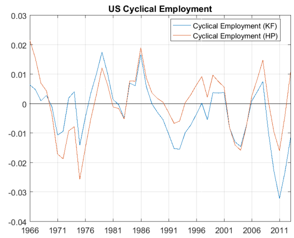
(a) 35-39 Actual and potential employment



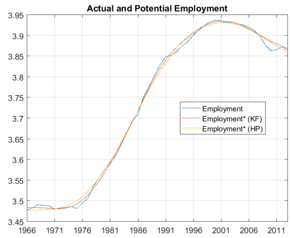
(b) 35-39 HP confidence bands



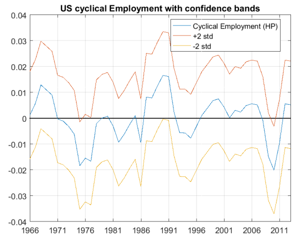
(c) 35-39 Actual and potential employment



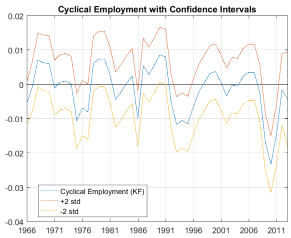
(d) 35-39 HP confidence bands



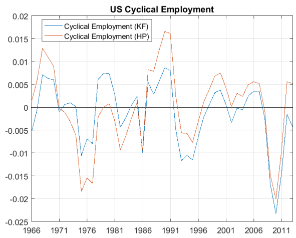
(e) 40-44 Actual and potential employment



(f) 40-44 HP confidence bands

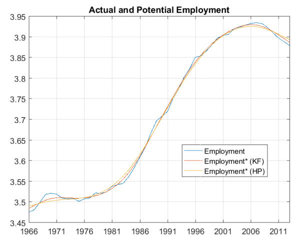


(g) 40-44 Actual and potential employment

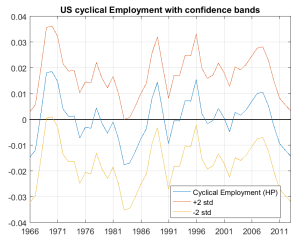


(h) 40-44 HP confidence bands

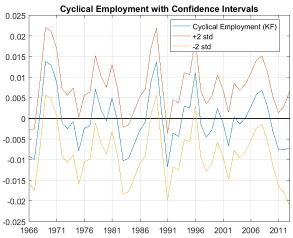
Figure 5.4: Results for females by age cohort 45-49 and 50-54



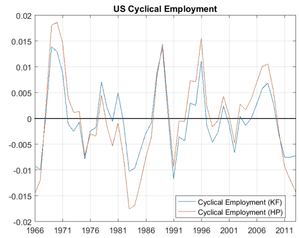
(a) 45-49 Actual and potential employment



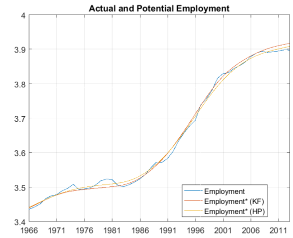
(b) 45-49 HP confidence bands



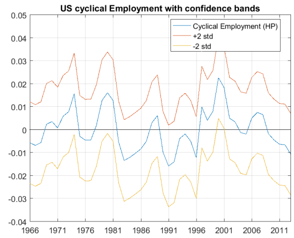
(c) 45-49 Actual and potential employment



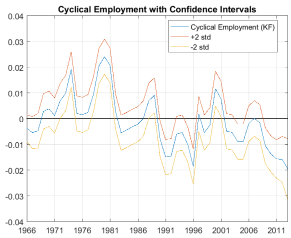
(d) 45-49 HP confidence bands



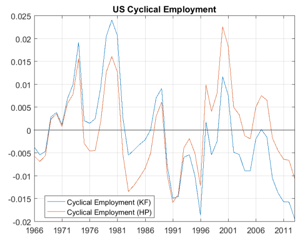
(e) 50-54 Actual and potential employment



(f) 50-54 HP confidence bands

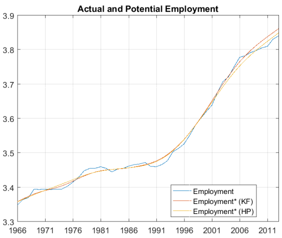


(g) 50-54 Actual and potential employment

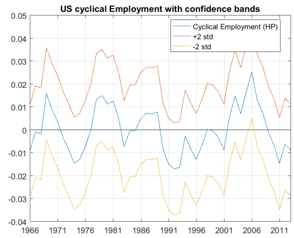


(h) 50-54 HP confidence bands

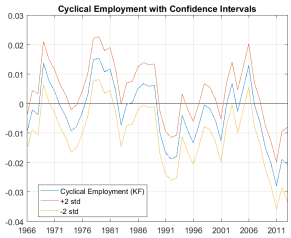
Figure 5.5: Results for females by age cohort 55-59 and 60-64



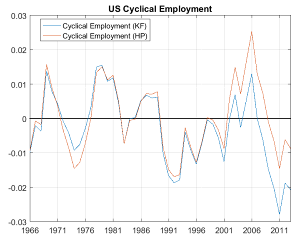
(a) 55-59 Actual and potential employment



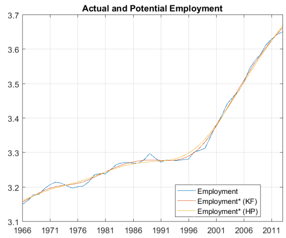
(b) 55-59 HP confidence bands



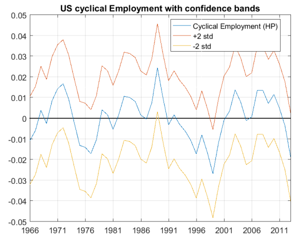
(c) 55-59 Actual and potential employment



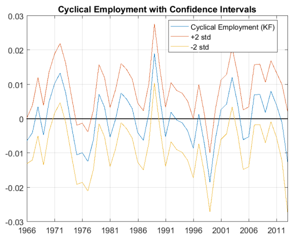
(d) 55-59 HP confidence bands



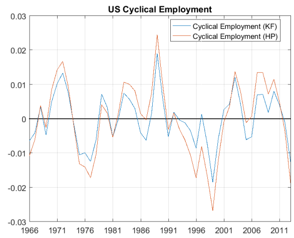
(e) 60-64 Actual and potential employment



(f) 60-64 HP confidence bands

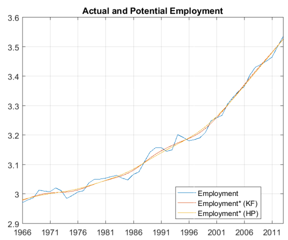


(g) 60-64 Actual and potential employment

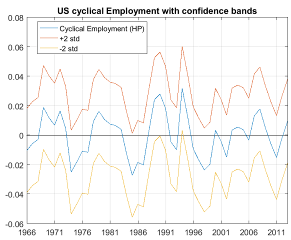


(h) 60-64 HP confidence bands

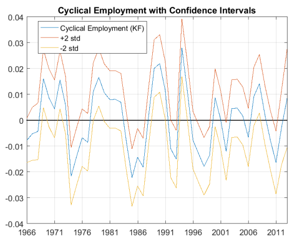
Figure 5.6: Results for females by age cohort 65+



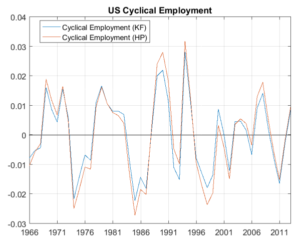
(a) 65+ Actual and potential employment



(b) 65+ HP confidence bands

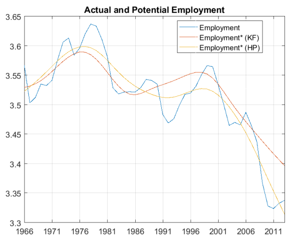


(c) 65+ Actual and potential employment

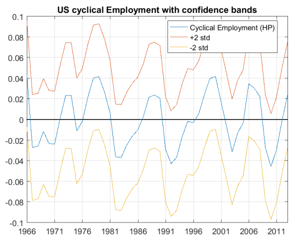


(d) 65+ HP confidence bands

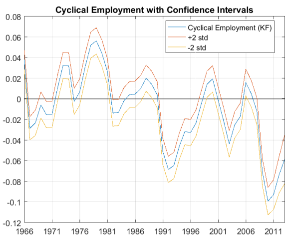
Figure 5.7: Results for males by age cohort 15-19 and 20-24



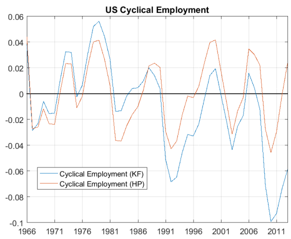
(a) 15-19 Actual and potential employment



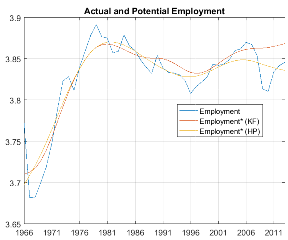
(b) 15-19 HP confidence bands



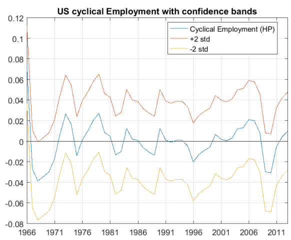
(c) 15-19 Actual and potential employment



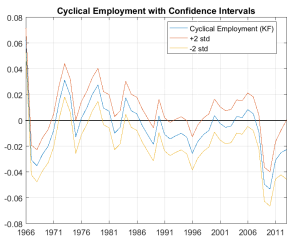
(d) 15-19 HP confidence bands



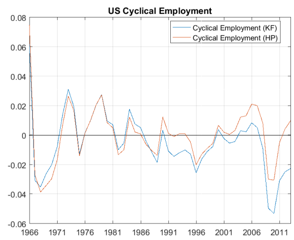
(e) 20-24 Actual and potential employment



(f) 20-24 HP confidence bands

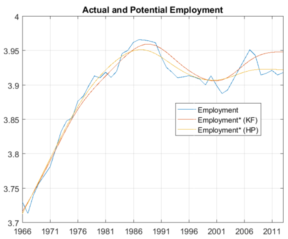


(g) 20-24 Actual and potential employment

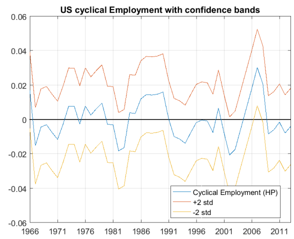


(h) 20-24 HP confidence bands

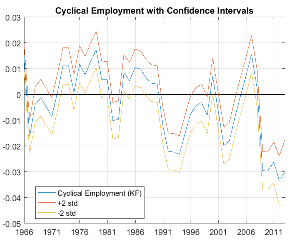
Figure 5.8: Results for males by age cohort 25-29 and 30-34



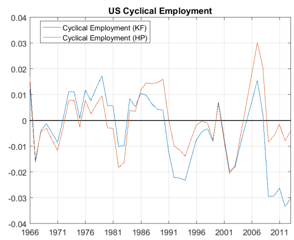
(a) 25-29 Actual and potential employment



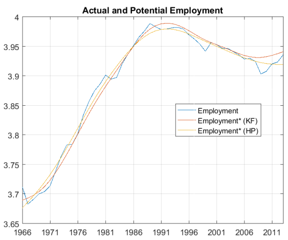
(b) 25-29 HP confidence bands



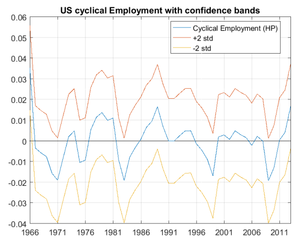
(c) 25-29 Actual and potential employment



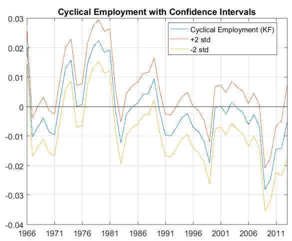
(d) 25-29 HP confidence bands



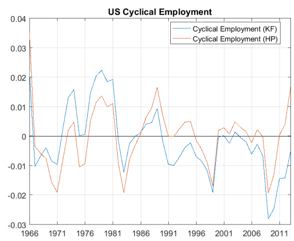
(e) 30-34 Actual and potential employment



(f) 30-34 HP confidence bands

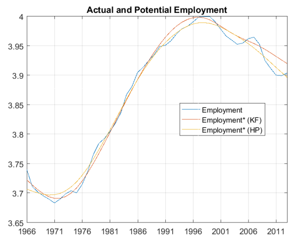


(g) 30-34 Actual and potential employment

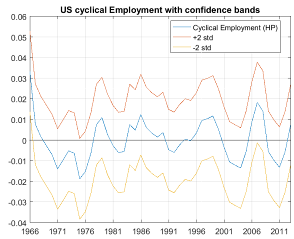


(h) 30-34 HP confidence bands

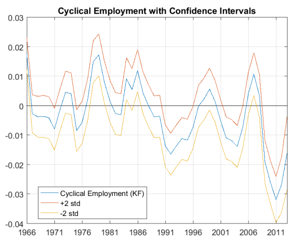
Figure 5.9: Results for males by age cohort 35-39 and 40-44



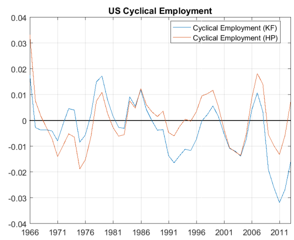
(a) 35-39 Actual and potential employment



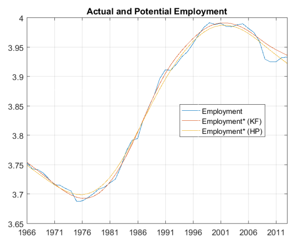
(b) 35-39 HP confidence bands



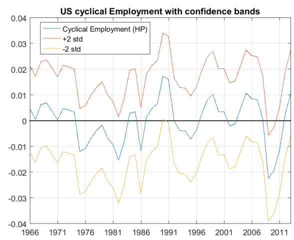
(c) 35-39 Actual and potential employment



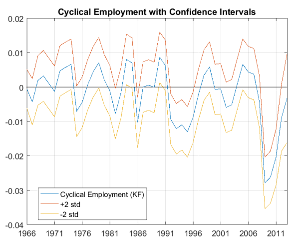
(d) 35-39 HP confidence bands



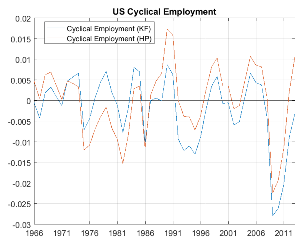
(e) 40-44 Actual and potential employment



(f) 40-44 HP confidence bands

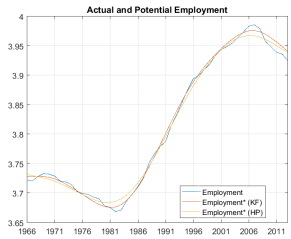


(g) 40-44 Actual and potential employment

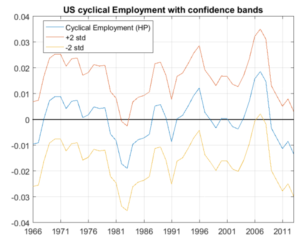


(h) 40-44 HP confidence bands

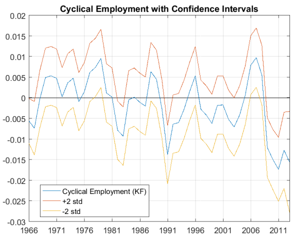
Figure 5.10: Results for males by age cohort 45-49 and 50-54



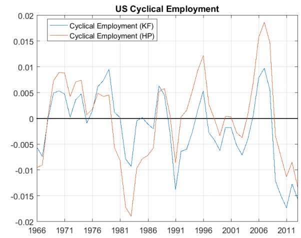
(a) 45-49 Actual and potential employment



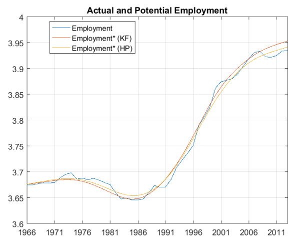
(b) 45-49 HP confidence bands



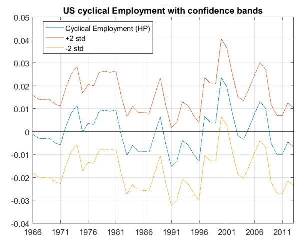
(c) 45-49 Actual and potential employment



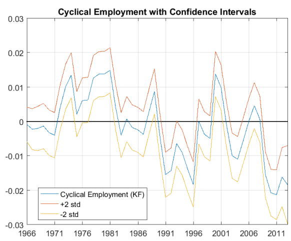
(d) 45-49 HP confidence bands



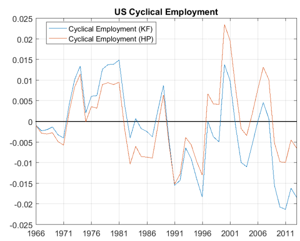
(e) 50-54 Actual and potential employment



(f) 50-54 HP confidence bands

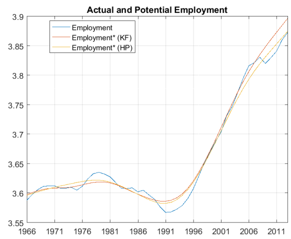


(g) 50-54 Actual and potential employment

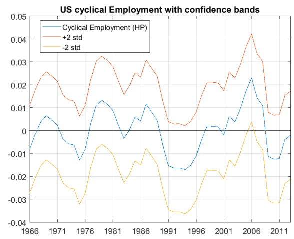


(h) 50-54 HP confidence bands

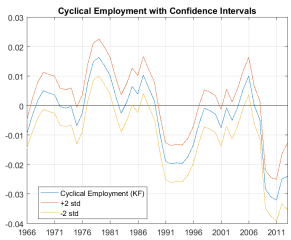
Figure 5.11: Results for males by age cohort 55-59 and 60-64



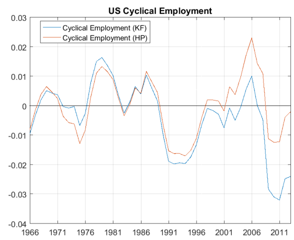
(a) 55-59 Actual and potential employment



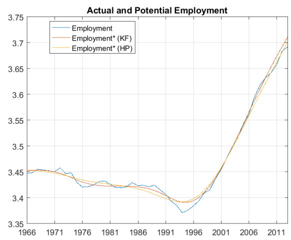
(b) 55-59 HP confidence bands



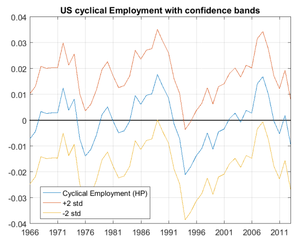
(c) 55-59 Actual and potential employment



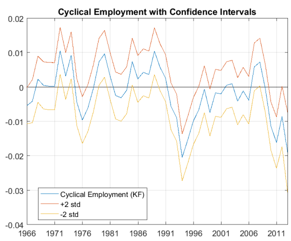
(d) 55-59 HP confidence bands



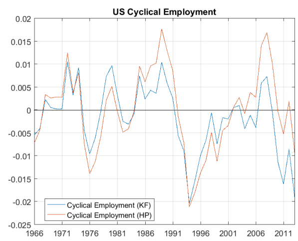
(e) 60-64 Actual and potential employment



(f) 60-64 HP confidence bands

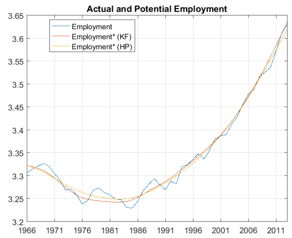


(g) 60-64 Actual and potential employment

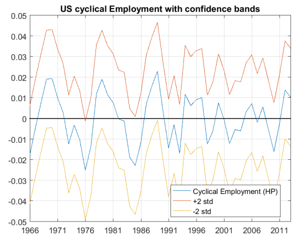


(h) 60-64 HP confidence bands

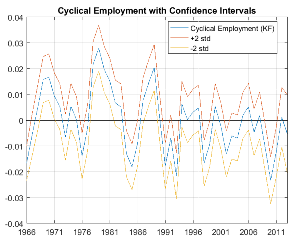
Figure 5.12: Results for males by age cohort 65+



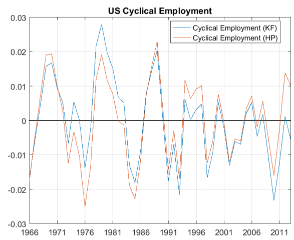
(a) 65+ Actual and potential employment



(b) 65+ HP confidence bands

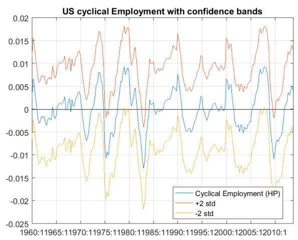


(c) 65+ Actual and potential employment

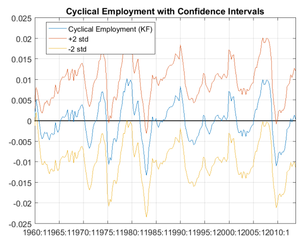


(d) 65+ HP confidence bands

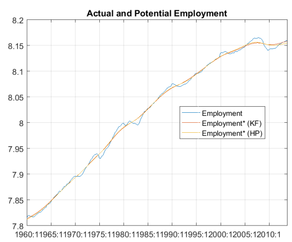
Figure 5.13: Quarterly figures, total employment



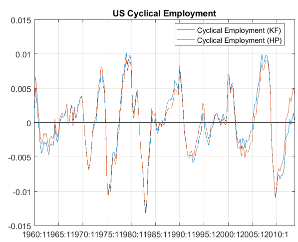
(a) Quarterly HP



(b) Quarterly KF



(c) Quarterly actual series



(d) Quarterly KF-HP